2016

Use of Protein Crosslinking and Tandem Mass Spectrometry to Study the PsbO, PsbP and PsbQ Extrinsic Proteins of Higher Plant Photosystem II

Manjula Priyadarshini Mummadisetti
Louisiana State University and Agricultural and Mechanical College, mmpdsetti@gmail.com

USE OF PROTEIN CROSSLINKING AND TANDEM MASS SPECTROMETRY TO STUDY THE PsbO, PsbP AND PsbQ EXTRINSIC PROTEINS OF HIGHER PLANT PHOTOSYSTEM II

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Biochemistry

by
Manjula Priyadarshini Mummadisetti
B.Sc., Kasturba Gandhi College, Osmania University, 2005
M.Sc., Osmania University, 2007
M.S., Louisiana Tech University, 2009
August 2016

ACKNOWLEDGEMENTS

First and foremost I want to thank my advisor Dr. Terry M. Bricker, who has been a tremendous mentor. He taught me, both consciously and unconsioucly, how to think logically from research perspective. I would like to thank you for encouraging my research and for allowing me to grow as a research scientist. Your advices, contributions of time, ideas and research funding through Department of Energy made my Ph.D. experience invaluable, productive and stimulating. The passion and enthusiasm you have for research was very contagious and you always inspired me even during the though times in pursuit of PhD. You have taught me that research is an art of logical thinking, fitting pieces of the puzzle together to form a functional plan and presentation of results to the scientific community. I am also thankful for having a successful biochemist as my thesis advisor. He helped to choose and develop the thesis topic and guided me over the years through the research project. I would like to express my sincere gratitude for his patience and constant guidance during my graduate program.

I would like to thank my thesis committee members, Dr. John Larkin, Dr. James Moroney, Dr. Michal Brylinski and Dr. Allison McFarland for serving on my thesis committee. I always enjoyed thoughtful discussions with Dr. Larkin, helping me see a project from different angles. I would like to thank you for your efforts and guidance on networking with people and for future career. Dr. Moroney has always been very enthusiastic to talk to me about my project or other scientific conversations. I would like to specifically thank Dr. Michal Brylinski for his constant guidance on protein modelling aspect of my project. My thesis wouldn't have been complete without his intellectual suggestions from his extensive expertize on protein model building. I would also like to thank Dr. William Wischusen and Dr. Christopher Gregg for providing me the

opportunity to explore and hone my teaching skills. I enjoyed learning this craft, while interacting with undergraduate students and trying to teach them in creative ways.

Additionally, I would like to thank Dr. Johnna Roose for being an amazing mentor. She helped me in learing the ropes of Plant Biochemistry and for being there to teach, inspire and push me to learn new techniques, and thinking of science from different perspectives. You have been more than a mentor to me, Johnna.

I would also like to thank our lab manager Ms. Laurie Frankel for her constant efforts in making this lab a wonderful place to work. Andrea, Annette, Adam, Nathalie and Kaitlyn are all undergraduates from this lab who helped in making this lab a most memorable learning and enjoyable experience.

A special thanks to my family members who supported me through out my higher education. Words cannot express how grateful I am to my mother (Mrs. Visalakshi), father (Mr. Kasinadh), sister (Sumalatha) and brother (Omkarnath) for all their sacrifices made on my behalf. This thesis is dedicated to my parents, their prayers helped me sustain this far. And finally, yet most importantly, I would like to thank my beloved husband (Mr. Sai Sharan Donthi) and his parents (Mr. Ramanath Donthi and Mrs. Yeshoda Ramanath) for being there for me and constantly supporting and encouraging me to stay focused towards the graduate program and my future career in research. I am especially thankful to Sai Sharan for always willing to be available, working around my schedule and helping me spiritually during the final stages of my graduate program.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

.mgf	Mascot Generic File
AFM	Atomic Force Microscopy
BS3	Bis (sulfosuccinimidyl) suberate
CAD	Collisionally Activated Dissociation
CID	Collision-Induced Dissociation
Cytb6f	Cytochrome b_6f
DSP	Dithiobis (succinimidyl propionate)
DSS	Disuccinimidyl Suberate
DTT	Dithiothreitol
ECD	Electron Capture Dissociation
EDC	1-ethyl-3-(3-diethylaminopropyl) carbodiimide
ESI	Electrospray Ionization
Fd	Ferredoxin
FDR.	False Discovery Rate
FNR	Ferredoxin: $NADP^+$ reductase
FT	Fourier Transform
HM	Homology Modelling
HTG	*n*-heptyl-β-D-thioglucoside
IAA	Iodoacetic acid
ICR	Ion Cyclotron Resonance
IEM	Inner Envelope Membrane
KFC	Knowledge based FADE and Contacts
LC	Liquid Chromatograph
LiDS	Lithium Dodecyl Sulfate
LMW	Low Molecular Weight
LTD	Lumen-Targeting Domain
MALDI	Matrix Assisted Laser Desorption Ionization
Metal Cluster	Mn_4CaO_5
MS	Mass Spectrometry
MSP	Manganese-Stabilizing Protein
NHS	N-hydroxysulfosuccinimide
NPQ	Non-Photochemical Quenching
OEC	Oxygen Evolving Complex

OEM	Outer Envelope of the Membrane
OMSSA	Open Mass Spectrometry Search Algorithm
PAGE	poly-acrylamide Gel Electrophoresis
PAR	Photosynthetically Active Region
PC	Plastocyanin
Pheo	Pheophytins
PSI	Photosystem I
PSI-BLAST	Position Specific Iterated–BLAST
PSII	Photosystem II
ROS	Reactive Oxygen Species
Sec	Secretory
SMN	Sucrose Mes NaCl
Sulfo-NHS	Sulfo-N-hydroxysulfosuccinimide
Tat	Twin Arginine Transport
TIC	Translocon at Inner membrane of Chloroplasts
TOC	Translocon at Outer membrane of Chloroplasts
TOF	Time-of-Flight
WOC	Water Oxidation Complex
α-DM	*n*-dodecyl-α-D-maltoside
β-DM	*n*-dodecyl-*β*-D-maltoside

ABSTRACT

Photosystem II (PSII) enzyme is a light-driven, water plastoquinone oxidoreductase present in all oxygenic photosynthetic organisms. The oxygen evolution process is catalyzed by the Mn_4CaO_5 cluster and an ensemble of intrinsic and extrinsic proteins which are associated with the photosystem. This metal cluster is stabilized and protected from exogenous reductants by the extrinsic proteins, PsbO, PsbP and PsbQ in higher plants, which are present on the lumenal face of PSII. No crystal structure for the higher plant PSII is currently available; consequently, the binding locations of these extrinsic proteins in PSII remain elusive. We have used chemical-crosslinkers Bis (sulfosuccinimidyl) suberate (BS3) and 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) to crosslink the extrinsic proteins in their bound state to PSII followed by identification of the crosslinked products by tandem mass spectrometry.

BS3 crosslinking identified the interacting domain of PsbP with PsbQ involving the PsbP residues ^{93}Y, ^{96}K and ^{97}T (located in the 17-residue loop 3A, ^{89}G-^{105}S) which are in close proximity ($\leq$11.4Å) to the N-terminal ^{1}E residue of PsbQ. We also found that this PsbP assumes a compact structure from the nine independent crosslinked residues between the N- and C-terminus of PsbP. This suggests that the N-terminus of PsbP, ^{1}A-^{11}K (which is not resolved in the current crystal structures), is closely associated with the C-terminal domain ^{170}K-^{186}A. Additionally, interacting domains of two PsbQ copies from different PSII monomers were identified. The residue pairs ^{98}K-^{133}Y and ^{101}K-^{133}Y of PsbQ were crosslinked. These residues are >30 Å apart when mapped onto the PsbQ crystal structure. Since BS3 can only crosslink residues which are within 11.4 Å, these residues are hypothesized as inter-molecular crosslinks of PsbQ.

Furthermore, EDC crosslinking provided structural information pertaining to the organization of the N-terminus, absent in the cyanobacterial-PsbO. In this study, twenty-four crosslinked residues located in the N-terminal, loop and the β-barrel region of PsbO were identified. The models incorporating crosslinking data suggests several differences in cyanobacterial- and higher plant-PsbO. The results on extrinsic proteins provide significant new information concerning the association of the extrinsic proteins with PSII and are valuable while proposing overall models of higher plant PSII.

CHAPTER 1. LITERATURE REVIEW

1.1 Oxygenic Photosynthesis

Oxygenic photosynthesis is an important process by which plants, algae and some bacteria convert light energy, carbon dioxide and water into carbohydrates and oxygen. There are also organisms like purple and green bacteria that carry out photosynthesis without producing oxygen. The process of photosynthesis in plants and algae occurs in specialized organelles called chloroplasts that contain an aqueous fluid called stroma, and a membranous system called thylakoids. The stacked thylakoids are called grana lamella, and unstacked thylakoids are called stroma lamellae. In cyanobacteria thylakoids are membrane staks of folded sheets present near plasma membrane (Frain et al., 2016). The thylakoid membranes contain huge membrane bound protein complexes that form supercomplexes which work in series as a unit. In oxygenic photosynthesis, also called linear electron transport, the complexes of Photosystem II (PSII), Cytochrome *b_6f* (Cyt *b_6f*), Photosystem I (PSI) and Ferridoxin (Fd)-$NADP^+$reductase (FNR) work in tandem to generate oxygen, reduce $NADP^+$ to NADPH, and generate a proton gradient in the lumenal space which is used to drive ATP synthesis by the chloroplast ATP synthase. These products are ultimately used up by the Calvin cycle for carbon fixation.

During the electron transfer reaction shown in FIGURE 1.1, electrons are withdrawn from water by water-splitting reactions at the PSII metal cluster and charge separation events within the D1/D2 proteins of PSII (described below) transfer electrons from Q_A to Q_B to form plastoquinol PQH_2. The plastoquinol then diffuses in the membrane to the Cyt *b_6f* complex, where it transfers one electron to a small copper containing protein called plastocyanin (PC) and the other electron to another PQ molecule within *b6f* complex. The PC after accepting an electron travels to the PS I complex and transfers its electron via ferredoxin (Fd), to the complex ferredoxin: $NADP^+$ reductase

(FNR) that reduces $NADP^+$ to NADPH in stroma. It has been hypothesized that Fd can also transfer electrons back to the *b6f* complex in one form of cyclic electron flow (Munekage et al., 2004).

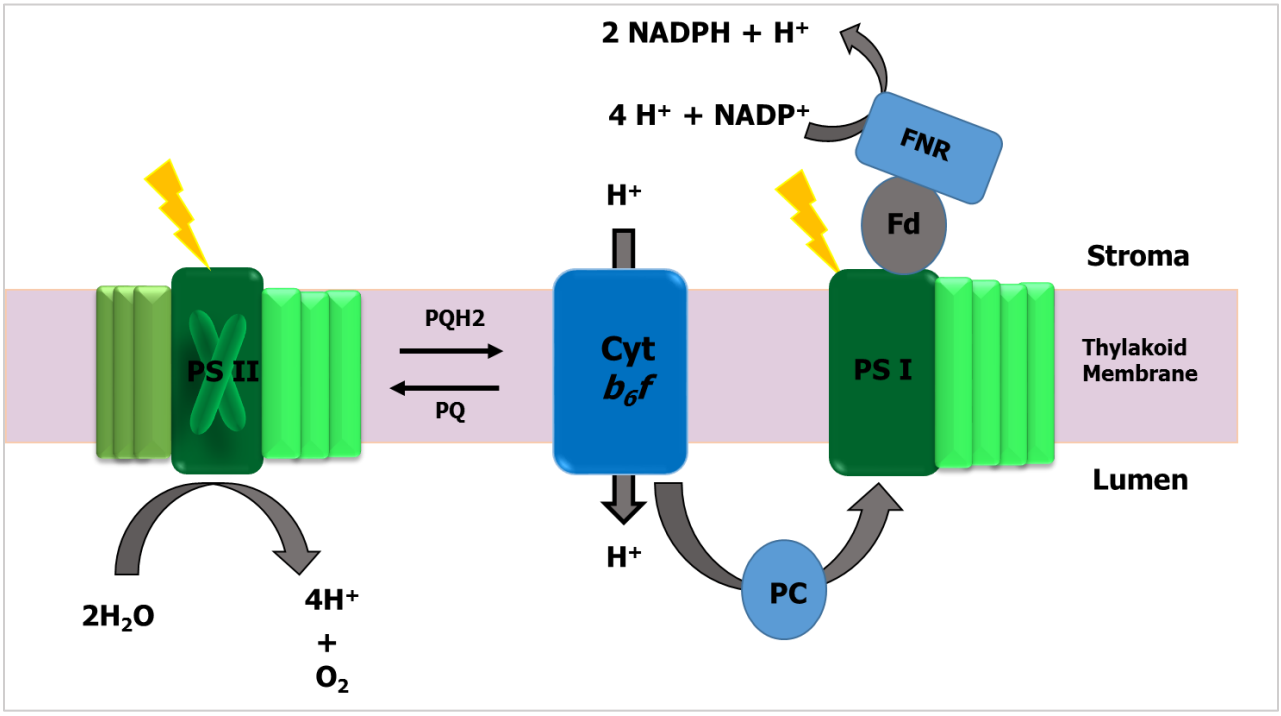


FIGURE 1.1. Linear Electron Transport of Photosynthesis.

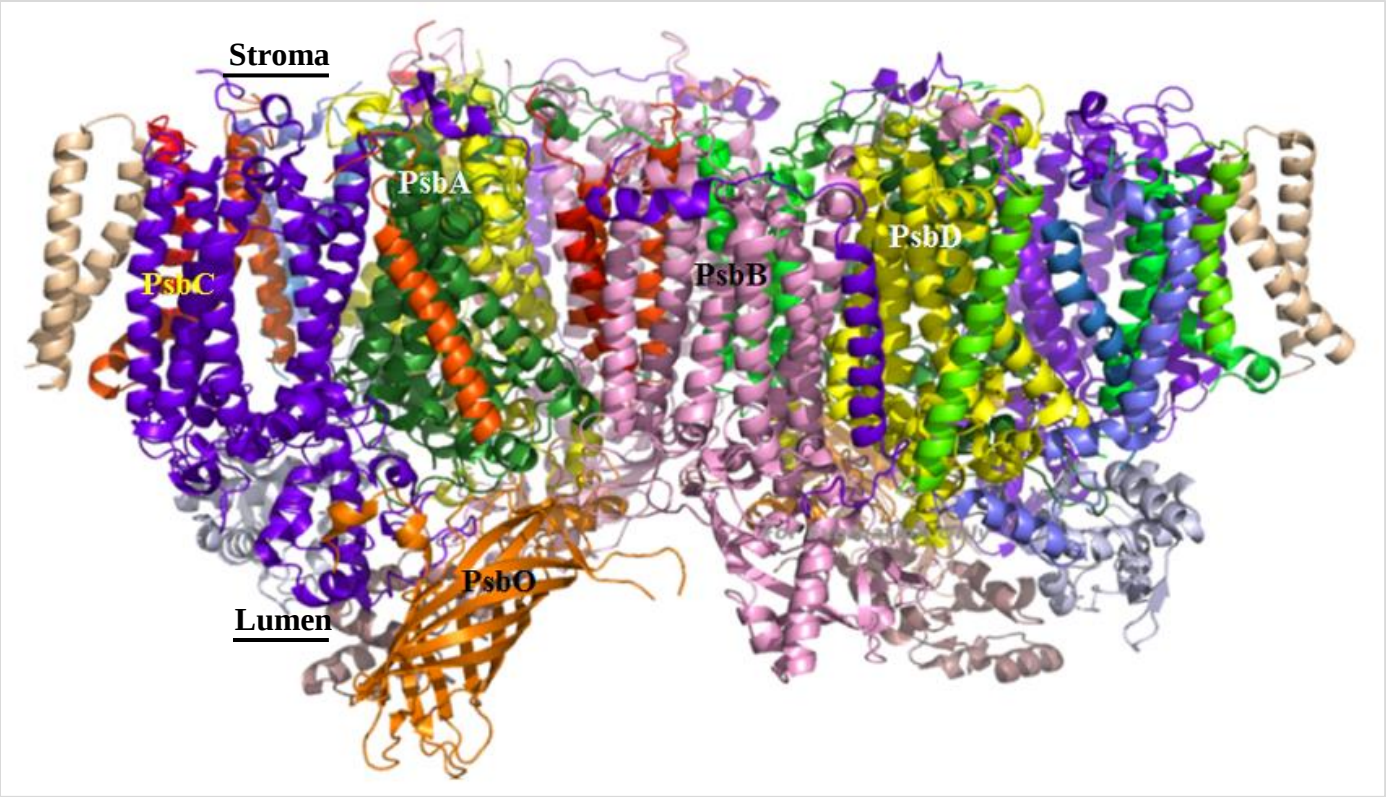
In this figure, protein complexes PSII, Cyt *b6f* and PSI are shown working in series. PSII and PSI are shown in green, *b6f* and FNR in blue. *b6f* acts as a proton pump. The proton gradient generated on the luminal side is used for ATP synthesis. The linear electron transfer reaction ultimately reduces $NADP^+$ to produce NADPH.

This photosynthetic electron transfer process is initiated by the capture of light by accessory pigments chlorophyll *a* (Chla), chlorophyll *b* (Chlb), carotenoids (like β carotene, lutein, violoxanthin etc.) and in cyanobacteria, phycobilins using the photosynthetically active region (PAR) of light in the 400-700 nm wavelengths range. When these pigments absorb a photon of light, they transfer the absorbed energy to Chl*a* to form a singlet excited state (^{1}chl*) and this state can transfer energy further downstream to the reaction center chlorophylls by resonance energy transfer. The excited chl can dissipate its energy in different ways, the primary use of this energy is for

photosynthesis (photochemistry), and excess energy is dissipated either as heat also known as non-photochemical quenching (NPQ) or it can be re-emitted as fluorescence. These three phenomena occur simultaneously (although in different proportions). In the following sections, PSII, with a focus on the extrinsic proteins from higher plants, will be discussed, along with their comparison to the cyanobacterial counterparts.

1.2 Photosystem II (PSII)

The architecture of PSII (FIGURE 1.2) has several intrinsic membrane proteins embedded in the thylakoid membranes, extrinsic proteins on the lumenal side, and an antenna system to capture light energy. PSII is a water-plastoquinone oxidoreductase that catalyzes the light-dependent oxidation of water into molecular oxygen. PSII exists as a 700 kDa homodimer. In all oxygenic organisms, light is trapped by and funneled through light-harvesting pigment protein complexes to the reaction center of PS II, where two of the core subunits, D1 and D2, located in the heart of PSII contain all the cofactors required for primary photochemistry. D1 and D2 are heterodimeric proteins that bind these cofactors arranged along the pseudosymmetric paths in PSII. These cofactors are two Chla molecules present as a dimer called P_{D1} and P_{D2}; two monomeric chlorophyll (Chla) molecules – Chl_{D1}, $Chl_{D2;}$ two pheophytins (Pheo) and two quinones (Q_A and Q_B). The FIGURE 1.3 illustrates the cofactors involved in electron transfer process within PSII. The charge separation primarily occurs on the D1 path, excitation of Chla forms the cation $[P_{D1}/P_{D2}]^{\bullet+}$ with a dominating state of $Chl_{D1}^{\bullet+}$ $Pheo_{D1}^{\bullet-}$ in almost all centers (Cardona et al., 2012; Saito et al., 2013a, b). This radical pair is stabilized by electron transfer to $Q_{A\,,}$ after which Tyr 161 on D1 also called TyrZ (Y_z) donates one electron to $P_{D1}^{\bullet+}$. And further $Y_z^{\bullet}$ oxidizes the Mn4CaO5 cluster, which then catalyzes the water splitting mechanism, to generate more electrons (Saito et al., 2013a). These electrons further stabilize the $Y_z^{\bullet}$ radical, and transfer electrons ultimately to the plastoquinone $Q_{A.}$



Umena et al., 2011

FIGURE 1.2. X-ray Crystal Structure of Photosystem II.

PSII dimer from T. vulcanus at 1.9Å resolution. Individual subunits are colored in different colors. PsbA (green), PsbB (light pink), PsbC (purple), PsbD (yellow) and PsbO (orange) are labelled. Cofactors associated with PSII are not shown.

Q_A is a single electron carrier, it transfers electrons to Q_B one at a time, and once Q_B is fully reduced by two electrons, it forms plastoquinol (QBH2). The plastoquinol at the Q_B site is loosely bound to the protein D1, and this can diffuse into the thylakoid membrane once it forms the QBH2 exchanging itself with another plastoquinone to bind at the Q_B site. The electrons from QBH2 are transferred to plastocyanin via Cyt b_6f (FIGURE 1.1). During this process of charge separation, several reactive oxygen species (ROS) are generated. ROS species can oxidize numerous surrounding residues in the PSII complex that impairs the structure and function of the complex.

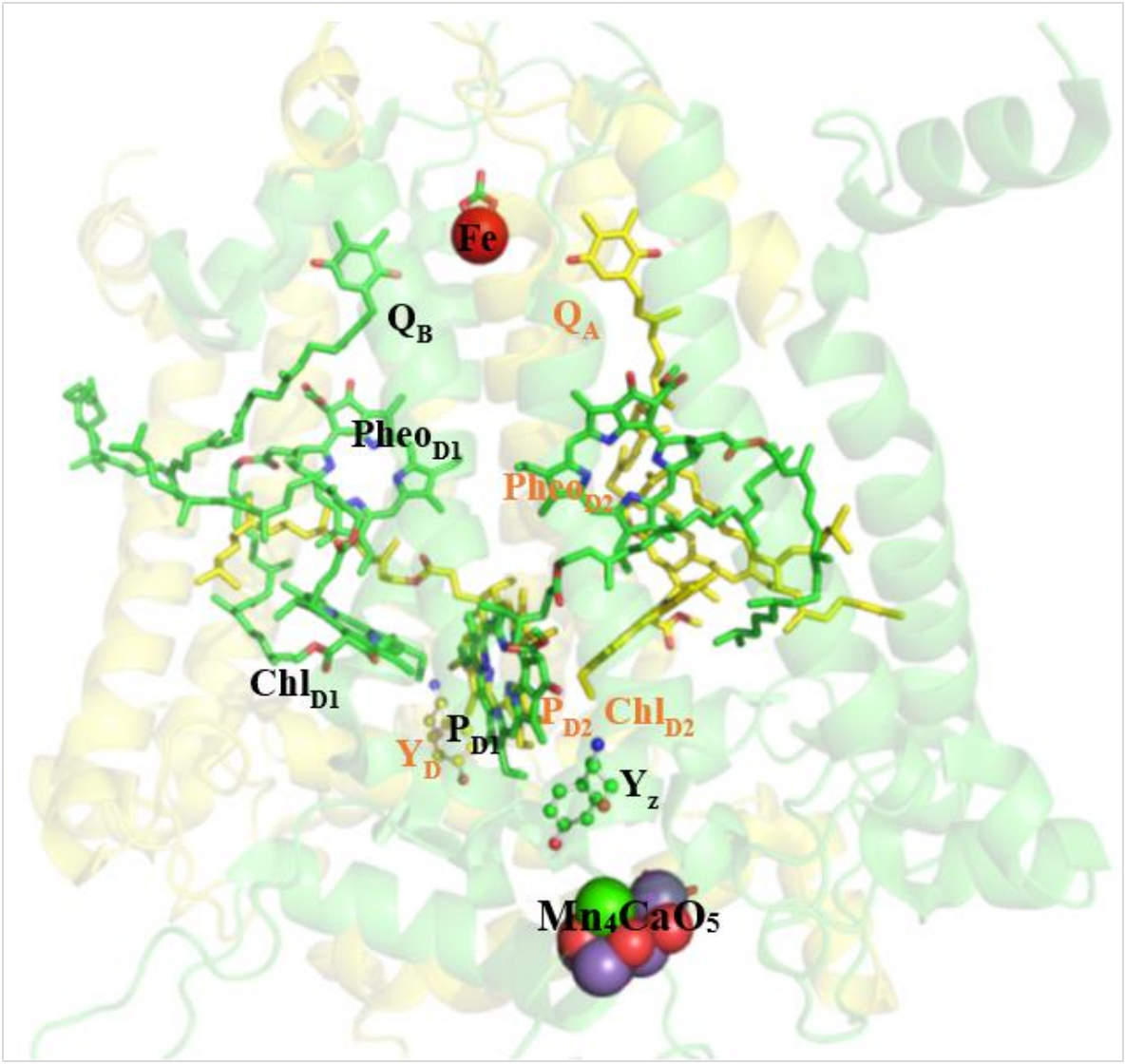


FIGURE 1.3. Cofactors in Electron Transfer Process of Photosystem II.

D1 protein and its cofactors are shown in green, and D2 protein along with the associated cofactors are shown in yellow. Redox-active tyrosines Y_z (present on D1) and Y_D (present on D2) are shown in ball and stick models. Chl_{D1}, P_{D1} and Chl_{D2}, and P_{D2} are chlorophyll pigments on D1 and D2 respectively.

Since PSII disassembly and reassembly is energetically an expensive process, there are several internal free-radicle quenchers which include both enzymatic and non-enzymatic antioxidants that help in scavenging these ROS species (Pospisil, 2014). These antioxidants can only theoretically scavenge and eliminate the ROS completely. When ROS are produced in excess, these

species are channeled from their site of production to the antioxidant catalytic center via specific channels (Frankel et al., 2012, 2013). This is considered as the first-line of ROS defense (Pospisil, 2014), the second line of defense is initiated outside the thylakoid membrane, once these ROS species are carried out of these channels. The putative ROS channels have been experimentally studied using radiolytic footprinting and tandem mass spectrometry by (Frankel et al., 2013).

In addition to ROS channels, there are other channels for the entry of the substrate water and the exit of the products, oxygen and protons, from the metal cluster. Potential channels for water, oxygen and protons have been proposed based on computational analysis using the program CAVER (Petrek et al., 2006), by looking at the continuous/contiguous spaces to accommodate water or molecular oxygen in PSII crystal structure available at that time (Gabdulkhakov et al., 2009; Ho and Styring, 2008; Murray and Barber, 2007). Similarly, to identify the regions of the protein around the metal cluster that are accessible to the bulk solvent (water and methanol in this case) were identified computationally using the cyanobacterial PSII crystal structure (Ho and Styring, 2008). The oxygen and water channels were further experimentally studied (Frankel et al., 2012, 2013) by observing the oxidation pattern of the potential channel residues. It was hypothesized that the water present in these channels (entering or exiting the metal cluster) would be susceptible to synchrotron radiation, and would generate hydroxyl radicles, which are capable of oxidizing the surrounding residues. The oxidized residues were thus studied by exposing the PSII membranes to synchrotron radiation at various time intervals (in seconds) and analyzed by tandem mass spectrometry. The authors identified several modified residues on D1, D2, CP47 and CP43, and compared them to the computational analysis published by the above-mentioned groups. Based on their analysis, 55% of the residues identified from D1, D2 and CP47 were similar to broad channel (Ho and Styring, 2008),

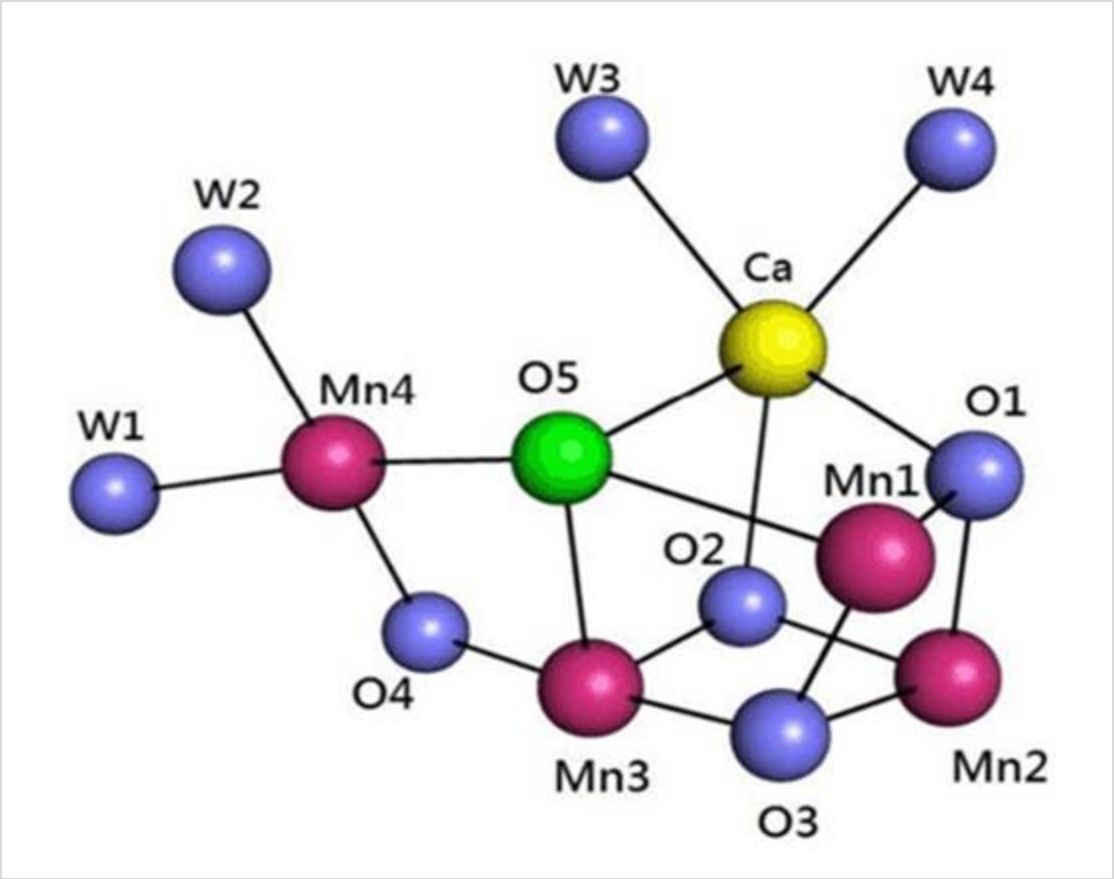
or channel (iii) and (c) suggested by other groups, and discussed in detail in recent reviews (Bricker et al., 2015; Ho, 2012).

PSII complexes have been crystallized in both monomeric (PDB ID: 3KZI, (Gabdulkhakov et al., 2009) and dimeric forms (Kamiya and Shen, 2003; Umena et al., 2011). The crystal structure with the best resolution has been studied from a thermophilic cyanobacteria at 1.9Aº in a dimeric form (Umena et al., 2011). Although the PSII dimer is considered the most active and functional form, functional monomeric complexes with similar oxygen evolving rates have also been reported (Kern et al., 2005; Shen, 1997) and more detailed discussion on monomeric complexes has been reviewed elsewhere (Watanabe et al., 2009). The PSII crystal structure from *Thermosynechococcus vulcanus* (henceforth *T. vulcanus*) illustrated in FIGURE 1.3 is composed of about seventeen membrane proteins, three extrinsic subunits, 35 chla, 11-12 al-trans β carotenes, one Mn_4CaO_5 cluster, one heme b, one heme c, two to three plastoquinones, two pheophytins, one non-heme iron, and about 25 lipids (Umena et al., 2011). Of the seventeen membrane proteins at least seven are highly essential for the photoautotrophic growth and oxygen evolution function of PSII. These seven proteins include D1, D2, CP43, CP47, Cytb559 α and β subunit and PsbL. The D1 and D2 proteins each consists of five transmembrane helices that contain all the contributing cofactors for the electron transfer and water splitting reaction. CP43 and CP47 are made of six transmembrane helices that bind to numerous chlorophyll molecules and serve as an internal antenna system. $Cytb_{559}$ is made of two subunits α (PsbE) and β (PsbF) and a heme group coordinated by two histidines contributed one from each subunit. PsbL is a low molecular weight (LMW) protein present at the monomer-monomer interphase of PSII along with two other proteins (Suorsa et al., 2004), PsbM and PsbT. Absence of PsbL causes decreased photoautotrophic growth, reduced number of PSII centers and impaired oxygen evolution (Luo et al., 2014). PSII containing these seven components

can evolve oxygen, however, only at minimal levels. Addition of very high, non-physiological levels of calcium and chloride causes proper oxygen evolution. The presence of extrinsic proteins (PsbO, PsbP and PsbQ in higher plants) helps in achieving optimal levels of oxygen evolution in the absence of high calcium and chloride. The extrinsic proteins include PsbO, PsbP, PsbQ and PsbR in higher plants (green lineage), and PsbO, PsbU, PsbV, CyanoQ and CyanoP in cyanobacteria.

The oxygen evolving complex (OEC) of PSII is made of an active site which is a chemically unstable complex, packaged during the PSII assembly process. The active site is made of Mn_4CaO_5 inorganic cluster buried on the lumenal side of D1 and CP43, ligated by waters and amino-acid side chains (Ferreira et al., 2004; Loll et al., 2005; Umena et al., 2011; Yano et al., 2006). The Mn_4CaO_5 cluster is protected by extrinsic proteins (PsbO, PsbP and PsbQ in higher plants and PsbO, PsbU and PsbV in cyanobacteria) and the lumenal domains of the intrinsic membrane proteins (PsbA, PsbB, PsbC and PsbD). These proteins help to shield and stabilize the cluster from the exogenous reductants.

The FIGURE 1.4 shows the arrangement of ions in the Mn_4CaO_5 inorganic cluster. It resembles a distorted cubane or a skewed chair and is based on the X-ray crystal structure of PSII (Umena et al., 2011) and quantum mechanics/molecular mechanics (QM/MM) calculations by Ishikita and Saito group. In the FIGURE 1.4, the corners of the cubane are occupied by three Mn ions (Mn1, Mn2, and Mn3), four oxygen atoms (O1, O2, O3 and O5 (shown in green)) and one Ca^{+2} ion. The Mn4, however, is located outside of the cubane, connected to the rest of the ions by μ-oxo-bridges formed by O4 and O5, and thus Mn4 is also known as dangler Mn. It is hypothesized that this distorted structure of the metal complex would help in structural rearrangements during the S-state transitions, and for efficient catalysis of water-splitting process.



Chu, 2013

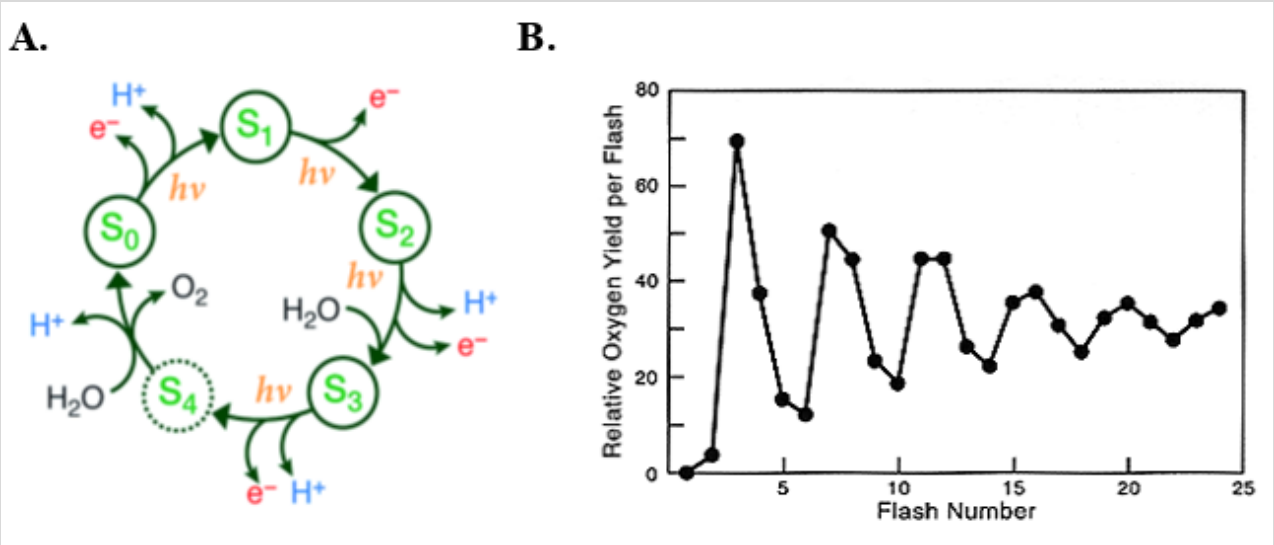
FIGURE 1.4. The Mn_4CaO_5 Cluster from the PSII Structure, 3WU2.

The X-ray crystal structure of PSII identified the Mn_4CaO_5 cluster organized as shown in this figure. Manganese (Mn) is shown as purple spheres occupying the corners of the cubane; oxygen (O) and waters (W) are shown in blue. Four water molecules are present around the Mn cluster of which, W1 and W2 are coordinated to Mn4, and W3 and W4 are coordinated to Ca.

The mechanism of oxygen evolution has been explained on the basis of the Bessel Kok model (Kok et al., 1970), and the discovery of the flash-induced period four oscillations was by Pierre Joliot (Joliot, 1969). This model has a lasting impact on the current models to date. Joliot observed that when a series of short flashes of light are provided to the green alga *Chlorella*, the oxygen evolution process oscillates in a series of four (FIGURE 1.5), with the maximum levels of oxygen being released at the third flash (Joliot, 1965, 1968, 2003). On continuing these series of flashes, the oxygen evolution dampens and subsequently reaches a steady-state level (Forbush, 1971; Joliot,

1969; Kok et al., 1970). The Kok model suggests that the PSII is activated by single turnover flashes leading to oxygen evolution, in which the water oxidation complex (WOC) cycles through five discrete redox states, also called S-states, named $S_0 - S_4$. The S-states advance by flashes of light. These states are characterized by their increasing order of oxidation, with S_0 being the most reduced and S_4 being the most oxidized. S_4 is also a transient state capable of oxidizing water to release oxygen, and hence oxygen is released during $S_3 \rightarrow [S_4] \rightarrow S_0$ transition (Han et al., 2012; McEvoy and Brudvig, 2006). The release of oxygen and binding of water molecule occur simultaneously (FIGURE 1.5 A).

S_1 is suggested as the stable and dominating state in the dark, with S_2 and S_3 being the metastable states that decay back to S_1 in a few minutes at room temperature, and S_0 oxidizes back to S_1 in a few minutes by $Y_D^{\bullet}$. TyrD (Y_D), PsbD:^{160}Y , is a redox active that can provide electrons to P_{D1}/P_{D2} chlorophyll, just like TyrZ. However, Y_D doesn't have a major effect on the charge transfer process. The replacement of this Tyr with Phe (Debus et al., 1988; Saito et al., 2013b) did not seem to affect the PSII function, thus the main function of Y_D known is to oxidize S_0 state to S_1. The dampening of oscillations in the Kok model were explained by misses and double hits. Misses in this model suggest the inability to advance to the next S-state. Double hits on the other hand represent a double turnover (moving two states ahead taking two photons of light) from one single flash. In terms of PSII, during the first flash P_{680}^{+} is generated, oxidizing the dark stable S_1 state to S_2. The second flash takes it to S_3 and the third flash to S_4, ultimately releasing oxygen at S_4, and returning back to S_0. The release of oxygen at the third flash is seen when S_1 state is predominant and successive cycles require four turnover flashes to extract four protons and electrons and ultimately release oxygen.



Krewald et al., 2015 life.illinois.edu

FIGURE 1.5. Oxygen Evolution Process Proposed by Kok, 1970.

A. the Figure shows the events occurring at each S-state. Oxygen evolution and entry of water occurs during S_3 to S_0 transition. The dotted circle around S_4 shows its transition state. With each flash of light, one electron and a proton are released. **B**. The figure illustrates oscillating pattern of oxygen yield with brief series of flashes.

1.3 The Extrinsic Proteins of Photosystem II in Higher Plants

The extrinsic proteins are present on the lumenal face of PSII and are required for enhanced oxygen evolution and optimal activity at physiological levels of calcium and chloride conditions (Bricker, 1992). Additionally, they protect the Mn_4CaO_5 cluster (henceforth, metal-cluster) from the exogenous reductants and reactants present in the surrounding aqueous phase. An array of extrinsic proteins binding to PSII differs with species suggesting their evolutionary changes and significance over time. The extrinsic proteins present in higher plants are PsbO, PsbP, PsbQ and PsbR. In cyanobacteria they are PsbO, PsbU, PsbV, CyanoP and CyanoQ. On the other hand, red algae have PsbO, PsbU, PsbV and PsbQ', while diatoms have Psb31 in addition to the proteins found in the red algae (Okumura et al., 2008). In this dissertation, only the proteins from higher plants will be discussed and compared to their cyanobacterial counterparts. Detailed reviews on the extrinsic

proteins can be found in (Bricker and Burnap, 2005; Bricker and Frankel, 2011; Bricker et al., 2012; Ifuku, 2015; Roose et al., 2016; Roose et al., 2007b).

In higher plants, these extrinsic proteins are encoded by the nuclear genes *psb*O, *psb*P, *psb*Q and *psb*R. The proteins are expressed as precursors that consist of a bi-partite (two consecutive) N-terminal signal sequences, one that targets the protein to chloroplast with the second targeting to lumen. The precursor proteins are guided to stroma in an unfolded state via Translocon at Outer membrane of Chloroplasts (TOC) complex located on the outer envelope of the membrane (OEM) and Translocon at Inner membrane of Chloroplasts (TIC) complex located on the inner envelope membrane (IEM). These translocons are GTPases, and require GTP and ATP to perform their transport function. After synthesis of the protein in the cytosol, the chloroplastic signal sequence is exposed for the transit of the proteins to stroma. Once these proteins reach the stroma, their chloroplast signal is cleaved off by the signal peptidase to expose the second transit peptide for lumenal proteins. The proteins extensively studied, traversing through these pathways, are PsbO, PsbP and PsbQ. The additional targeting sequences of these lumenal proteins are called Lumen-targeting Domain (LTD). LTD containing proteins are generally soluble proteins. A subset of these proteins is targeted to the lumen by the Secretory (Sec) pathway, and others by Twin Arginine Transport (Tat) pathway. PsbO is an example of Sec pathway, where the proteins are transported in their unfolded state and the process is driven by ATP hydrolysis. PsbP, PsbQ and their homologs on the other hand are translocated into the lumen by the Tat pathway. These Tat pathway substrates are in their folded state, with an "Arg-Arg" (twin Arginine) motif that is recognized by the Tat pathway receptor, and uses proton motive force (*pmf*) across the membrane to derive energy for the translocation process. The Tat is the only known pathway that uses *pmf* instead of NTPs for driving translocation of proteins to their destination.

The Structure, Function and Location of PsbO in PSII

PsbO is a key extrinsic protein present in all organisms containing PSII. Calcium and chloride are the essential cofactors of oxygen-evolving complex. PsbO, also known as Manganese-stabilizing protein (MSP), supports the Mn_4CaO_5 cluster with the optimal levels of Cl^- and protects it from the exogenous reductants (Popelkova et al., 2008). Additionally PsbO, along with PsbP and PsbQ, may assist in maintaining optimal levels of Ca^{+2} as well.

PsbO is highly conserved among higher plants, algal and cyanobacterial species. Sequence alignments from various species indicate that this protein is highly conserved with 43% of the residues being either entirely conserved or conservatively replaced among different species (Bricker and Burnap, 2007). Spinach PsbO exhibits 47% identity and 65% sequence similarity with cyanobacterial protein. The only available high resolution structure of PsbO is from the cyanobacterial PSII of *Thermosynechococcus vulcanus* (PDB ID: 3ARC) at 1.9 Å resolution (Umena et al., 2011). The red algal structure from *Cyanidium Caldarium* has been resolved at 2.77 Å resolution (Ago et al., 2016). Since the red algal structure of PsbO has several unresolved regions, only cyanobacterial structure is used here for our comparisons to understand spinach PsbO in the following sections.

Structure of the PsbO Protein

Structurally, PsbO has a body made of eight antiparallel beta strands arranged to form a barrel or cylindrical shape, illustrated as a ribbon diagram in FIGURE 1.6. The axis of the PsbO barrel is oriented at a 40º angle to the plane of the PSII core. Additionally, this structure has two helices and several loop regions. In this figure, the *T. vulcanus* structure, PDB ID 3ARC (Umena et al., 2011), is shown in orange, and the spinach PsbO sequence was threaded on the program Swiss-

model, using the cyanobacterial structure as a template. These two structures of PsbO are aligned to each other for proper comparison.

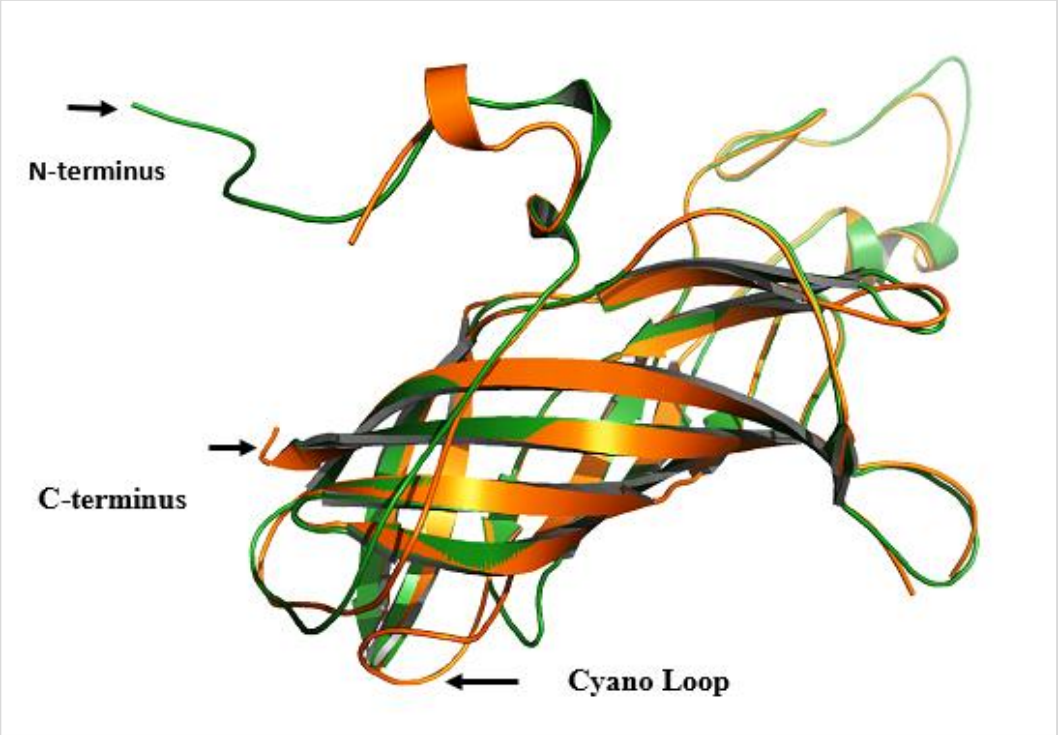


FIGURE 1.6. Cyanobacterial PsbO Vs Spinach PsbO.

Cartoon in orange is cyanobacterial PsbO, and green is spinach PsbO threaded using cyanobacterial structure as a template, using the program Swiss-model.

As seen in the figure, these structures are overall very similar, however, they do exhibit some prominent dissimilarities in their loop regions. Two major alterations between these structures are at the N-terminus and at the cyano loop region. Spinach PsbO has an extended N-terminus, 12 residues longer than the cyanobacterial PsbO, as illustrated in the figure. The Cyano-loop is an extended loop region from ^{129}T to ^{142}F residues; the analogous loop in spinach is much smaller in the region ^{136}G-^{141}F. It is suggested that these differences somehow cancel out each other's effect, and can function equally well in their respective organisms. PsbO subunits contains two cysteines, reported to form a disulfide bridge, which is conserved in all organisms. These loops and beta barrel

regions are suggested to interact with the intrinsic membrane proteins via charge pair interactions. The potential salt bridge and cation - ∏ interactions have been studied and reviewed earlier, (Bricker et al., 2012) using programs KFC (knowledge based FADE and contacts) server (Darnell et al., 2008) and CaPTURE (Gallivan and Dougherty, 1999) respectively, of which salt-bridge interactions are illustrated in TABLE 1.1. Structurally, this protein along with other extrinsic proteins PsbU and PsbV in cyanobacteria act as a lumenal cap to the Mn_4CaO_5-cluster protecting it from the exogenous reductants. The FIGURE 1.7 illustrates the location PsbO with respect to the metal cluster.

TABLE 1.1. Putative Salt-Bridge Interactions of PsbO in *T.vulcanus*.

Salt bridge interactions of PsbO identified using KFC and cation-π interactions by CaPTURE. The residues in parenthesis refer to the putative inter – monomer interactions. The numbering scheme is similar as found in the PS II structure 3ARC.

PsbO:CP 43	PsbO: CP 47	PsbO:D1	PsbO:D2
Charge-pair interactions			
^{8}D:^{362}R, 3.05 Å	^{169}D:^{385}R, 4.88 Å	^{69}K:^{103}D, 3.2 Å	^{160}K:^{302}E, 3.41 Å
^{99}D:^{381}K, 2.45 **Å**	^{169}D:^{422}R, 2.70 Å	^{73}R:^{104}E, 3.39 Å	^{228}H:^{310}E, 3.19 Å
	^{179}E:^{422}R, 4.27 Å		
	^{179}E:^{423}K, 5.20 Å		
	(^{59}K:^{307}E, 4.21 Å)		
Cation-π interaction			
^{7}Y:^{370}R, 4.67 Å			

Bricker et al., 2012

Function of the PsbO Protein

Removal of PsbO has different effects depending on the type of organism and whether it was examined *in vivo* or *in vitro*. *In vivo* mutant studies from cyanobacteria and *Chlamydomonas reinhardtii* (henceforth *Chlamydomonas*) suggested severe effects on the organisms' oxygen evolution ability. *Synechocystis* sp. PCC 6803 (henceforth *Synechocystis* 6803) mutants specifically,

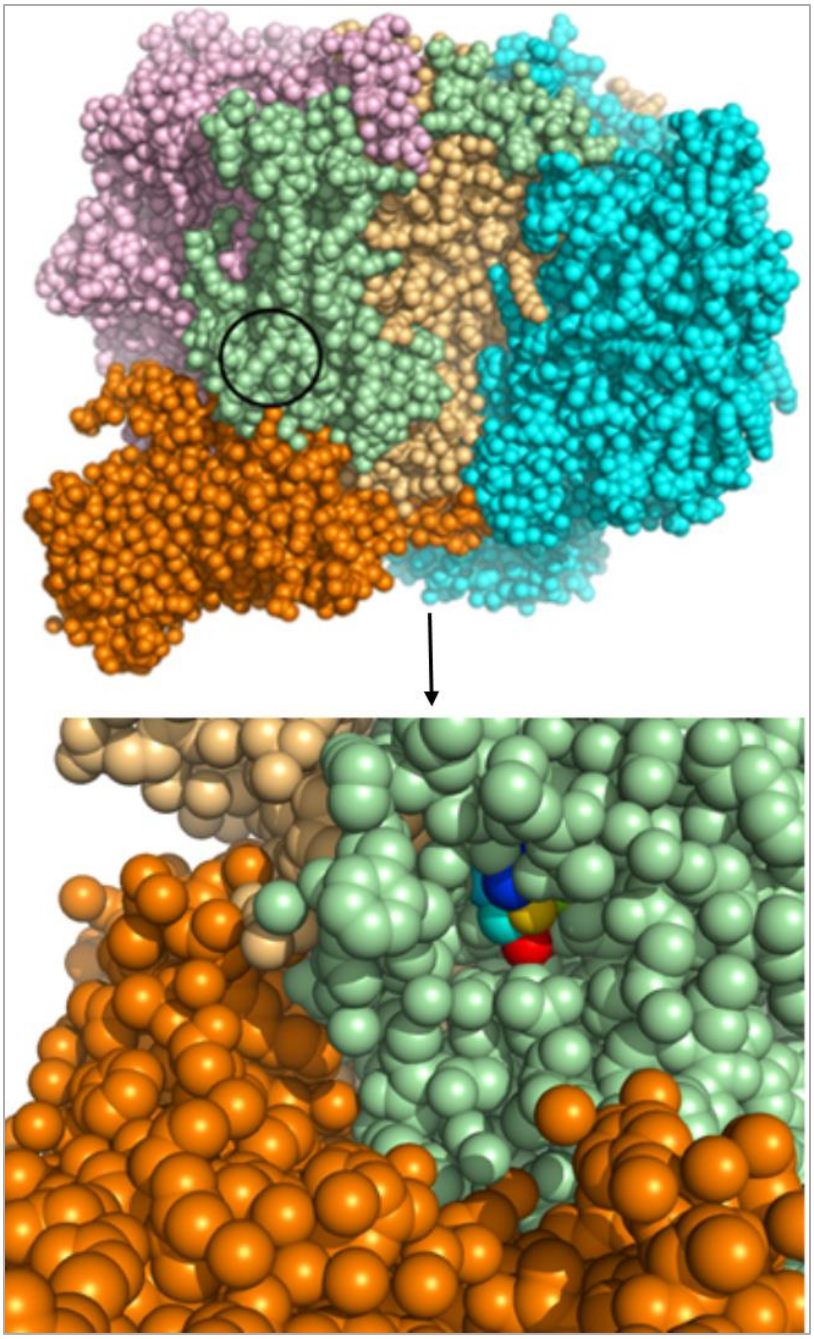


FIGURE 1.7. PsbO, a Lumenal Protein Cap for Mn-cluster.

A. The surface rendering of PsbO is shown in orange color, PsbA (D1) in green, PsbD (D2) in yellow, PsbB (CP47) in pink, PsbC (CP43) in blue. **B**. FIGURE A is rotated 180º anticlockwise, and zoomed in to see Mn-cluster, shown in rainbow colored spheres.

in the absence of PsbO, could still evolve oxygen at rates of about 30% of the wild type (Burnap and Sherman, 1991). It was hypothesized later that the *in vivo* levels of Ca^{+2} and Cl^{-} were likely sufficient to photo ligate the lost Mn ions and hence, help in the oxygen evolution process (Bricker, 1992).However, in the *Chlamydomonas* mutant FuD44, severe phenotypic effects were seen in the absence of PsbO. This mutant couldn't grow photoautotrophically, couldn't assemble PSII and was ultimately unable to evolve oxygen (Mayfield et al., 1987a). Similarly in Arabidopsis, P*sbO-1* mutants showed low photosynthetic yield and couldn't grow photoautotrophically (Murakami et al., 2002, 2005).

In vitro experiments on intact PSII membranes from spinach have been very useful in understanding the importance of PsbO in oxygen evolution. Spinach, an interesting tool and a representative of higher plant systems, has been used in numerous studies on the extrinsic proteins by release-reconstitution experiments. Removal of PsbO from spinach PSII membranes by $CaCl_2$ treatment leads to destabilization of the Mn_4CaO_5 cluster and, under low chloride (< 0.1 M) conditions, results in uncoupling of two of the four Mn ions from the metal-cluster (Mavankal et al., 1986). In presence of high Ca^{+2} (>0.1 M), the $CaCl_2$-washed PSII membranes, which lacked PsbO, could still evolve oxygen but at lower rates (20-40%) compared to the control. The NaCl-urea washed membranes (with their PsbO released) could evolve oxygen only in the presence of high levels of both Ca^{+2} and Cl^{-} (Bricker, 1992). These findings indicate an important role of PsbO in ion-retention and in stabilizing the Mn-cluster for the oxygen evolution process.

PsbO is very strongly bound to PSII by electrostatic interactions (first and second stoichiometric binding constants are K_{a1} 0.42 nM^{-1} and K_{a1} of o.67 nM^{-1} respectively, probably from two copies of PsbO (Leuschner and Bricker, 1996)) and can only be extracted in presence of high salt concentrations. It has been observed that 1-2M NaCl is not sufficient to extract PsbO from PSII

membranes, and this instead requires a higher charge density of divalent ions like Ca^{+2} (e.g. treatment with 1M $CaCl_2$) to disrupt its interactions with PSII. Additionally, it can also be extracted at high pH (e.g. 1M Tris) or in presence of a chaotropic agent (like urea). In presence of urea (2.6M), generally a low concentration of NaCl (200mM) is sufficient to disrupt the hydrogen bonding interactions between PsbO and PSII. The treatment with Tris also extracts the Mn ions from the metal cluster, and hence reconstitution of PsbO back to PSII doesn't restore full oxygen evolution activity (Leuschner and Bricker, 1996; Miyao, 1984, 1987).

Earlier studies on spinach PSII membranes depleted of PsbO suggested its role in S-state transition of Mn-cluster. The thermo-luminescence measurements indicated that the loss of PsbO caused slower transition of $S_3 \rightarrow S_0$ (Taka-aki Ono, 1985). Similar results were obtained from flash oxygen yield experiments (Miyao, 1987), which indicated that these membranes (lacking PsbO) exhibited extended S_2 and S_3 states, consequently causing slower turnover of $S_2 \rightarrow S_3$ and $S_3 \rightarrow [S_4] \rightarrow S_0$ transitions. Similar findings were also identified in studies from *Arabidopsis thaliana* and *Synechocystis 6803* mutants lacking PsbO1 (the isoform required for normal PSII activity) and PsbO respectively (Burnap et al., 1992; Imre Vass, 1992; Liu. et al., 2007).

To understand the contribution of PsbO to PSII, specifically at the site of Mn_4CaO_5, studies using Fourier transform infrared spectroscopy (FTIR) and isotope labelling have become very useful, these studies suggest changes in PsbO during state transitions in the metal-cluster (Offenbacher et al., 2013). These studies used globally labelled ^{13}C-PsbO, reconstituted onto unlabeled, highly active and fresh PSII core preparations. FTIR analysis suggested conformational changes in Mn-cluster upon removal of PsbO during $S_1 \rightarrow S_2$ state transition (using light minus dark spectra), and structural conformation changes within PsbO from isotope edited spectra (^{12}C-PsbO minus ^{13}C-PsbO).

PsbO structure and the arrangement of its residues in its tertiary structure, adds more light to its association to PSII. PsbO surface is particularly rich in aspartate and glutamate (D/E) residues, forming clusters, and several of these clusters are associated with water molecules (Bondar and Dau, 2012; Shutova et al., 2007) that are resolved in the PSII crystal structure. The crystal structure of PSII has been extremely useful in some of the computational studies as well. These studies were performed to understand the water dynamics and hydrogen-bonding network within PSII. It is suggested that the carboxylates of these D/E clusters participate in extended hydrogen bonding network. This network is hypothesized to help in inter-subunit interaction, protein conformational changes for its stability and movement of water and H^+ acting as a channel to their exit pathway (Lorch et al., 2015).

PsbO Copy Number in PSII Monomer

One important question, which still remains unanswered, is the copy number of PsbO per PSII. The X-ray crystal structure of cyanobacterial PSII has one copy of PsbO, however the biochemical analysis indicates the presence of two copies of PsbO per monomer in higher plants. The stoichiometry of this protein has been studied using various experiments like binding-reconstitution, mutation analysis and protein quantification (for detailed reviews, see Bricker and Burnap, 2005; Bricker and Frankel, 2011; Bricker et al., 2012; Ifuku, 2015; Roose et al., 2016; Roose et al., 2007b; Zubrzycki et al., 1998).

Extraction of PsbO by butanol/water phase partition suggested the presence of at least 1.7 moles of PsbO tightly bound per mole of PSII (Yamamoto et al., 1987). Experiments using protein quantification, and verification of the extinction coefficient of PsbO, suggested two copies of PsbO per PSII monomer (Xu and Bricker, 1992). Some other studies used various buffers to extract PsbO along with Mn (Leuschner and Bricker, 1996). These authors tested the rebinding of PsbO to the

PSII membranes containing, no Mn bound (1M Tris treated PSII membranes), two Mn bound to PSII (membranes treated with 1M $CaCl_2$ and 5mM NaCl (low Cl^-)) and all four Mn bound (membranes treated with 1M $CaCl_2$ with 200mM NaCl) to the Mn_4CaO_5 the cluster. It was observed that irrespective of the number of Mn ions associated, the number of PsbOs that reconstituted back to PSII remained unaltered, although, the binding affinities (k_d) of PsbO proteins are different in each case. Also, the binding curves showed a positive cooperativity in these experiments (by definition, cooperativity requires more than one binding component) with a Hill coefficient of about 1.6 - 2.2. Single particle analysis (Kuhl et al., 1999) on PSII core particles extracted from *Synechococcus elongates* suggested the presence of extra density which was suggested to be possibly due to the second copy of PsbO, but required further analysis for direct evidence.

The structural variation in the cyanobacterial and higher plant (spinach) PsbO (described above) led to studying the extended N-terminus in spinach. Several years of work on N-terminal truncation mutants suggested the presence of two binding domains in spinach, which are highly conserved across higher plants and green algae. These domains include ^{7}T-^{10}E and ^{15}T-^{16}Y regions (Popelkova et al., 2002b) of which, the region spanning ^{7}T-^{10}E is absent in cyanobacteria (Popelkova et al., 2003). Further analysis suggested that these regions are highly essential in higher plants, and deletion of either of these domains led to reduced reconstitution of PsbO (40% of the control) to PSII (Popelkova et al., 2002a). Threonine at ^{7}T and ^{15}T are particularly interesting and are highly required for the binding functionality of two PsbO subunits (Popelka and Yocum, 2012). Other regions in the N-terminus of PsbO, ^{1}E-^{6}L domain along with the ^{8}Y-^{14}K and ^{16}Y-^{18}E domains are not required for the functional binding of one of two copies of PsbO to PSII, however, these domains are required to prevent the nonspecific binding of PsbO (Popelkova et al., 2002b). Additionally,

higher fluorescence measurements and oxygen evolution assays on PsbO mutants with N-terminal truncations have also suggested that two copies of PsbO are required (Roose et al., 2011b).

1.4 PsbP and PsbQ Structures, Function and Location in PSII

PsbP and PsbQ are the other important members of higher plants PSII lumenal proteins and are predicted to have evolved from CyanoP and CyanoQ of cyanobacteria. These apparently underwent several genetic and functional modifications during evolution to the current higher plant forms. PsbP and PsbQ are structurally very similar to CyanoP and CyanoQ respectively (Balsera et al., 2005; Cao et al., 2015; Jackson et al., 2012; Michoux et al., 2014) as shown in FIGURES 1.8 and 1.9 respectively, however, functionally they resemble PsbU and PsbV of cyanobacteria with respect to protecting the Mn4CaO5 cluster from the exogenous reductants and in maintaining optimal levels of Ca^{+2} and Cl^{-} (Bricker, 1992; Ljungberg et al., 1983).

CyanoP and CyanoQ

The X-ray crystal structure of cyanobacterial PSII from *T. vulcanus* provides the exact location of PsbO, PsbU and PsbV, however it lacks these two proteins, CyanoP and CyanoQ, probably because of their low abundance and/or loss of protein during PSII isolation. The function of CyanoP is not known, because of its low abundance in the CyanoP-PSII complexes in biochemical preparations (Ishikawa et al., 2005; Thornton et al., 2004). The mutants lacking PSII did not contain CyanoP, suggesting that CyanoP associates with PSII, or PSII is required for CyanoP accumulation (Ishikawa et al., 2005). The isolation of CyanoP-containing complexes was reported by Ifuku's group, (Aoi et al., 2014), who tried to isolate complexes associating with CyanoP using different detergents. Three detergents (*n*-dodecyl-*β*-D- maltoside (*β*-DM), *n*-dodecyl-*α*-D-maltoside (*α*-DM) and *n*-heptyl-*β*-D-thioglucoside (HTG)) were used in this study.

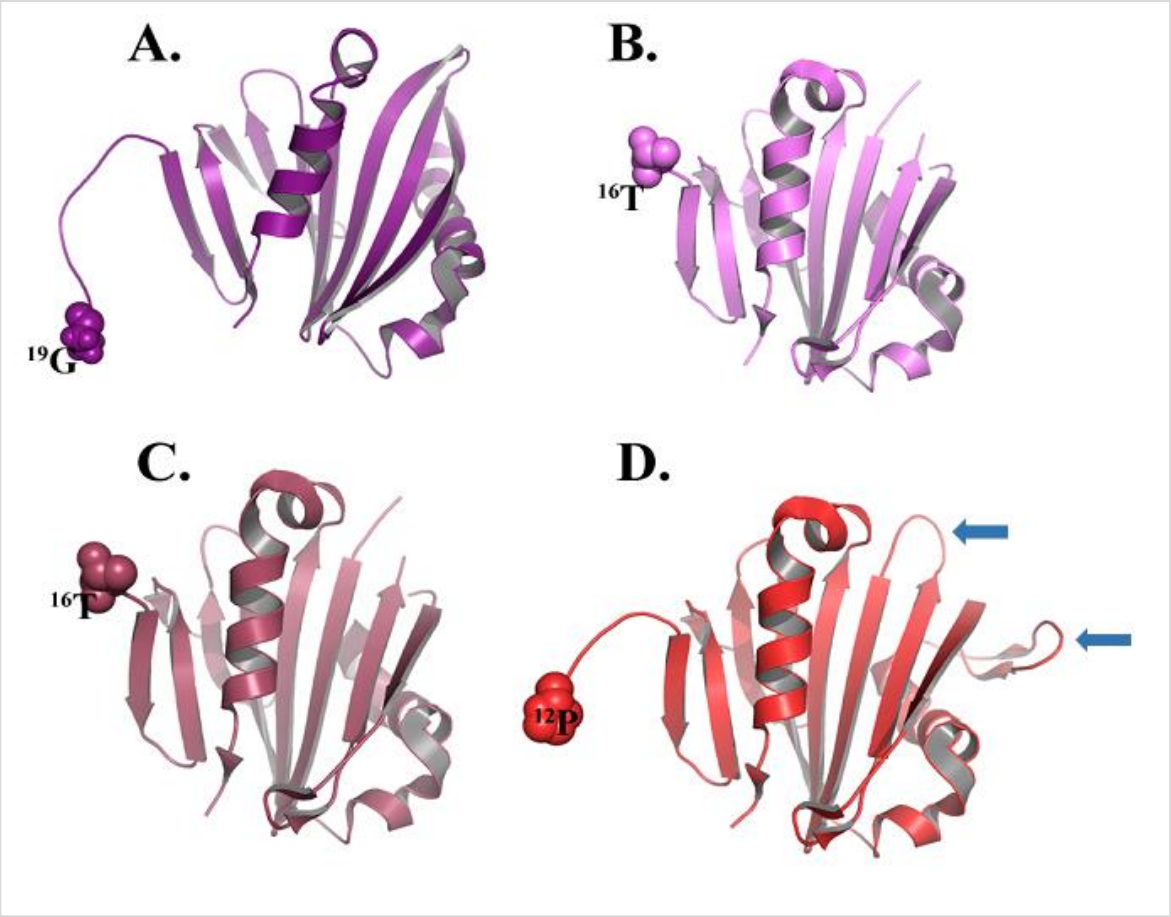


FIGURE 1.8. The PsbP Protein from Different Organisms.

The figure shows very small structural changes among higher plant PsbP proteins, however much larger changes from CyanoP. **A**. CyanoP from cyanobacteria (PDB: 2LNJ). **B**. PsbP from tobacco (PDB: 1V2B). **C**. PsbP from spinach, Zn bound (PDB: 2VU4). **D**. PsbP from spinach, Mn bound (PDB: 4RTI), the arrows shown in blue shows the additional regions resolved in this structure, with the binding of Mn ions.

The authors identified that only α-DM and HTG solubilized-membranes contained CyanoP attached to PSII, and soluble fraction obtained during the extraction of the PSII membranes also lacked free CyanoP. Additionally, α-DM (0.7%) solubilized membranes had only a minor fraction of CyanoP associated with PSII, while the majority existed as free CyanoP, suggesting a need for better purification methods. A deletion mutant of CyanoP revealed its functional importance. Although the mutant cells were able to grow photoautotrophically, subtle yet measurable changes were observed

in the deletion mutant under $CaCl_2$-depleted conditions (Aoi et al., 2014). 77K fluorescence emission spectra suggested an energy transfer anomaly from phycobilisomes to reaction centers under -$CaCl_2$ conditions and even under normal BG-11 conditions on the third day subculture of the mutant cells.

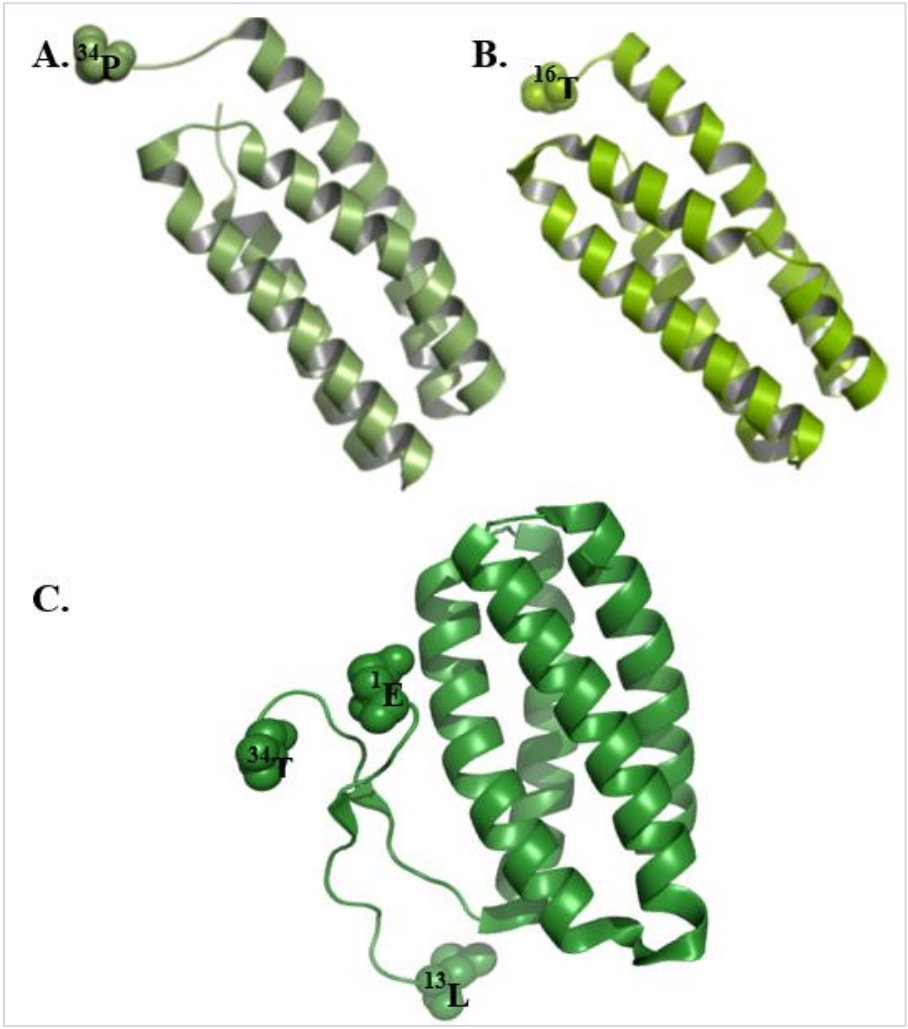


FIGURE 1.9. The PsbQ Protein from Different Organisms.

A. X-ray crystal structure of CyanoQ at 1.6 Å resolution from *Thermosynechococcus elongatus* (PDB ID: 3ZSU, (Michoux et al., 2013)). **B.** CyanoQ at 1.8 Å resolution from *Synechocystis* (PDB ID: 3LS0, (Jackson et al., 2010)). C. PsbQ at 1.49 Å resolution from spinach (PDB ID: 1VYK, (Balsera et al., 2005)).

Since CyanoP is a lumenal protein and phycobilisomes are the antenna systems present on the cytoplasmic side of the thylakoid membrane, CyanoP may be responsible for the long distance, transmembrane transfer of energy through conformational changes (Roose et al., 2016).Similar long-distance defects were observed upon removal of PsbP and PsbQ (Roose et al., 2010) in spinach PSII (slow electron transfer from Q_A^- to Q_B) and in ΔPsbU mutant (PsbU is one of extrinsic proteins in cyanobacteria) in S*ynechocystis* (Veerman et al., 2005), where defects in energy transfer from phycobilisomes to reaction center and electron transfer within PSII were observed. Based on the identified effects of CyanoP in PSII, and its low abundance, it was suggested that perhaps CyanoP is an assembly factor (Aoi et al., 2014; Thornton et al., 2004). Much recently, however, surface plasmon resonance spectroscopy identified the location of CyanoP in *T. vulcanus*. This protein occupied similar location as PsbO in PSII, and was only identified in the PSII assembly intermediates lacking the extrinsic proteins (PsbO, PsbU and PsbV). In this experiment, the amount of CyanoP reconstituted to PSII dimer was much less than the inactive, monomeric PSII from *T.vulcanus*. This was the first study to directly indicate the role of CyanoP in PSII assembly (Cormann et al., 2014). Additionally, the double mutants, *ΔcyanoPΔycf48*, were compared to the individual mutants and the wild type strain in *Synechocystis* (Jackson and Eaton-Rye, 2015). The *ΔcyanoP* mutant did not have any significant phenotype, and the *ΔcyanoPΔycf48* mutants were similar to *Δycf48* mutant except during PSII assembly. During PSII assembly, the number of assembled complexes identified in the double mutants was greater than in the *Δycf48* single mutant. This suggested that CyanoP potentially perturbs the PSII assembling capability still remaining in the absence of Ycf48.

As mentioned earlier, CyanoQ and PsbQ share structural resemblance in having a four-helix up-down bundle (PDB ID: 2LNJ, (Jackson et al., 2012) and PDB ID: 3ZSU, (Michoux et al., 2014)). The N-terminus of CyanoQ however, is lipidated, which is not seen in PsbQ. The sequence identity

of CyanoQ with the spinach PsbQ is only 17%. CyanoQ is not found in any of the currently available PSII crystal structures. This protein was suggested as a stoichiometric unit associating to PSII (Thornton et al., 2004). A six-histidine tag on the C-terminus of CyanoQ, aloowed extraction of the entire PSII complex by affinity purification confirming its association to PSII (Roose et al., 2007a). When CyanoQ-His-PSII complexes were compared to the CP47-His-PSII complexes (PSII extracted with a His-tag on CP47), the CyanoQ-His-PSII were found highly active compared to the CP47-His-PSII complexes. CP47-complexes were overall less active, probably because CP47-tag could isolate a heterogenous mixture of PSII centers, with PSII in their monomeric or dimeric states and any intermediary state undergoing assembly or repair (during assembly/repair, the extrinsic proteins are added much later in the process).

The deletion of *cyanoQ* (Δ*cyanoQ*) in *Synechocystis* could not affect the normal growth and photoautotrophy (Thornton et al., 2004). These mutants, however, evolved less oxygen than the wild type cells in the absence of $CaCl_2$. Additionally, these mutants were unable to grow photoautotrophically under Ca^{+2}, Cl^- and Fe depleted conditions, suggesting the role of CyanoQ in high stress conditions. Another interesting finding was that these deletion mutants could evolve more oxygen in presence of higher $CaCl_2$ (20mM) than with 5mM $CaCl_2$ (Kashino et al., 2006; Thornton et al., 2004). Additionally, these mutants were also sensitive to hydroxylamine (a reducing agent) than the wild type, suggesting an important role of CyanoQ on the extrinsic side of PSII in protecting the oxygen evolving complex (Kashino et al., 2006).

The hypothesis that one copy of CyanoQ is stoichiometrically bound per monomer of PSII (Liu et al., 2014; Roose et al., 2007a; Thornton et al., 2004) was revisited recently. Histidine-tagged CyanoQ could isolate PSII complexes, however, at higher concentration of imidazole, the extracted PSII complexes contained four copies of CyanoQ (instead of one) and instead lacked PsbU and PsbV

proteins (Liu et al., 2015). A model was presented explaining this result, which suggested binding of multiple copies of CyanoQ during PSII assembly process. It was further hypothesized that some of these copies are dissociated from PSII, along with addition of PsbU and PsbV during the maturation of PSII complex.

Crosslinking studies in *Synechocystis* identified crosslinked-residues within CyanoQ, which would only be possible if these residues were from two copies of CyanoQ (Liu et al., 2014). In these studies, the authors proposed the two CyanoQ copies were from two different monomers and hence they placed the CyanoQ in the PSII dimer interface. The N-terminus of CyanoQ is short (compared to spinach PsbQ) and has a lipid moiety at its N-terminus (which helps its association with the membrane). Hence, this protein would have to be very close to the thylakoid membrane while associating with PSII along with the rest of the extrinsic proteins (Roose et al., 2016).

PsbP Family of Proteins

The PsbP family of proteins constitutes CyanoP and its structural homologs in higher plants. The higher plant PsbP-family consists of the protein PsbP (described more in detail below), which is an extrinsic subunit of PSII (PsbP1 and PsbP2 pseudo-gene in Arabidopsis (Ifuku et al., 2010)), two PsbP-like proteins (PPL1 and PPL2) and seven PsbP-domain proteins named PPD1-7 (Bricker et al., 2013). All the members of the PsbP family of proteins were identified in lumenal proteome, extracted from the soluble fraction of thylakoid proteins (Peltier et al., 2002; Schubert et al., 2002). These proteins have evolutionarily diverged in their functions, but retain structural similarity. Homologous sequences have similar structures, however this is not directly related to their function, since small changes in structure can result in major change in function. Sequence alignments for PsbP family of proteins have been reviewed in detail (Bricker et al., 2013). Their structures contain several differences, with PPL and PPD proteins showing their characteristic insertion or deletion of certain

domains or extensions at their N-or C-termini. The functional characterization of PPL proteins in *Arabidopsis thaliana* (henceforth *Arabidopsis*) suggests that PPL1 acts as a PSII assembly/stability factor and is required for efficient repair of photo damaged PSII (Ifuku et al., 2010; Ishihara et al., 2007). Structurally and functionally, PPL1 is more similar to CyanoP from cyanobacteria than PsbP. PPL2 may have role in the NDH complex since it was found to be co-expressed with NDH complex subunits and is required for their accumulation (Ishihara et al., 2007).

Sequence alignments and phylogenomic studies of the PsbP-domain or PPD proteins indicated that the majority are specific to green lineage which includes the green alga *Chlamydomonas*, the moss *Physcomitrella patens* and other land plants (Ishihara et al., 2007; Sato, 2010). The functions of these proteins are not well understood. PPD1 is required for PSI assembly, specifically at the lumenal face of PSI and for the turnover of PsaA and PsbB proteins (Liu et al., 2012; Roose et al., 2014). PPD2 is categorized in NDH-related group (Ifuku et al., 2008), found in singlet-oxygen dependent signaling in *Chlamydomonas* (Brzezowski et al., 2012). PPD5 is a stress responsive signaling protein, also involved in strigolactone synthesis (Roose et al., 2011a). And PPD6 is also categorized in the stress responsive group and is a putative target of thioredoxin (Hall et al., 2010). These proteins have been reviewed in detail elsewhere (Bricker et al., 2013).

Higher Plant PsbP in PSII

Higher plant PsbP and PsbQ, involved in stabilizing the Mn_4CaO_5 cluster, binds to PSII by electrostatic interactions. These proteins when removed from PSII by 1M NaCl, the oxygen evolving ability of PSII is dropped to 25% of the control, and this effect is reversed upon rebinding of these proteins (Ljungberg et al., 1983; Miyao and Murata, 1983, 1985; Miyao, 1984). PSII can evolve oxygen in the absence of these two proteins only at high concentrations and non-physiological levels of Ca^{+2} and Cl^{-} (Ghanotakis et al., 1984a; Miyao and Murata, 1984; Nakatani, 1984). Hence, PsbP

and PsbQ are required to optimize and lower the requirement of Ca^{+2} and Cl^- ions around the Mn_4CaO_5 cluster. The mechanism underlying cofactor optimization is still a puzzle.

The N-terminus of PsbP is hypothesized to be critical for the ion retention function and binding to PSII. Measurements of oxygen evolution indicated that N-terminally truncated PsbP Δ15-PsbP and Δ19-PsbP were unable to reconstitute onto the PsbP depleted PSII, and hence couldn't optimize the Ca^{+2} and Cl^- requirement necessary for oxygen evolution (Ifuku et al., 2005a). This was explained in case of Δ19-PsbP with the hypothesis that the truncation led to instability in its structure, which was further supported by the access of its domains to trypsin digestion, not seen in wild type PsbP. It can now be deduced from the current crystal structure 4RTI that Δ19-PsbP was unstable because the residues ^{12}P-^{17}E form a random coil and residues ^{18}F-^{22}N form a β-sheet with the β-strand ^{26}F- ^{31}P. A deletion in any region of the β-sheet (as seen in Δ19-PsbP) would break the hydrogen bonding interactions and destabilize the structure, and hence making it accessible to proteolysis. Previous crosslinking studies report that the N-terminal ^{1}A of PsbP was crosslinked to ^{57}E of the α-subunit of Cyt b_{559} (Ido et al., 2012), a core subunit of PSII. Additionally, the N-terminus of PsbP can modulate the redox potential of Cyt b_{559} (Nishimura et al., 2016). Consequently the importance of the N-terminus is evident. Thus Δ15-PsbP was unable to restore the ionic requirement because it lacked the domain interacting with the PSII membrane proteins, in addition to maintaining ionic concentration at the oxygen-evolving site.

The current high-resolution crystal structure of higher plant PsbP is available from spinach at 1.8Å resolution (PDB: 4RTI, (Cao et al., 2015)). This structure contains two manganese ions, Mn1 and Mn2. Mn1 is strongly and stably bound to PsbP in a pocket of positively charged residues and coordinated to the side chains of ^{144}H and ^{165}D and a Cl^- ion as seen in the FIGURE 1.10. PsbP has a αβα structure, with three α helices and ten β strands. The previous structure of PsbP (PDB:

2VU4, (Kopecky et al., 2012)) was bound to Zn, instead of Mn, and had three unresolved regions. Two of the three regions and some portion of the third region (N-terminus) are resolved in the 4RTI structure. And yet, the N-terminus of PsbP, highly important for PSII binding and ion retention, from residues ^{1}A- ^{11}K is not determined in any of the crystal structures available so far.

FTIR spectroscopy is a popular technique to probe conformational changes in a protein or a protein complex. The FTIR measures vibrations of the atoms in a compound. The vibrations that lead to net change in the dipole moment appear as peaks in this spectroscopy. Hence this method has been used to identify changes in the structure of proteins by measuring the vibrational changes in the amide bond. The amide I band represents the vibrations associated with the C=O stretching (peak near 1650 cm^{-1} for alpha-helix and 1630-1640 cm^{-1} for beta-sheet) and amide II band arises mainly due to N-H bending (peak near 1550 cm^{-1}) (Barth 2007). The vibrations from amide bonds are similar for interacting protein, and hence isotope-edited infrared spectroscopy has become useful. Isotope labelling allows a shift in these amide bands, thus allowing a greater different in the vibration spectra. Similar studies (isotope-edited IR) have been performed to identify conformational changes around the Mn_4CaO_5 cluster using ^{13}C labelled PsbP. This ^{13}C labelling helped in identifying the changes specifically arising from PSII. In this experiment, no conformational changes were observed in the PsbP itself (Ono et al., 1992; Tomita et al., 2009). Additionally, FTIR difference spectroscopy also detected conformational changes in $S_2Q_A^-$/ S_1Q_A upon removal of PsbP, and these changes were restored on reconstitution of intact PsbP (Ido et al., 2012). The spectra indicated that there were no changes in the ligands from the Mn_4CaO_5 cluster. PsbP and PsbQ extrinsic proteins apparently work together in the ion retention process around the Mn_4CaO_5 cluster. Contrary to wild type PsbP, Δ15-PsbP could not restore the conformational changes in PSII when reconstituted onto the PSII lacking PsbP and PsbQ. Addition of PsbQ was able to restore these changes. Additionally Ifuku group

(Kakiuchi et al., 2012) discovered that the binding of Δ15-PsbP (which does not bind PSII) to PSII along with PsbQ was able to restore the oxygen evolution function to about 50% of the wild type. The residue ^{144}H of PsbP was reported to be essential for the ion retention ability of PsbP and this residue is highly conserved among PsbP and CyanoP (Kakiuchi et al., 2012).

In spite of several studies on extrinsic proteins, the structure of the N-terminus of PsbP remains unknown. The ambiguity in the binding location of extrinsic proteins has given rise to several studies like transmission electron microscopy (on intact PSII membranes), negative staining of PSII membranes from Arabidopsis and spinach, cryoelectron tomography of chloroplasts and grana stacks from spinach. In a separate study to understand the protrusion of the extrinsic proteins from the plane of the membrane, Atomic Force Microscopy (AFM) was performed (Phuthong et al., 2015). This identified the small and large protruding regions of PSII-OEC. However, in-depth analysis is necessary to understand the exact location of the extrinsic proteins in higher plant PSII.

The function of PsbP in Chlamydomonas has been examined in two mutants, BF25 and FUD39, both of which lack this protein (de Vitry et al., 1989; Mayfield et al., 1987b). These mutants could accumulate the reaction center proteins (CP43, CP47, D1 and D2) only to a level of 65% and 30% respectively. The BF25 mutant could grow photoautotrophically albeit, very slowly, and evolve oxygen at very low rates (about 5% of the wild type). It was also observed that in these mutants, PsbO and PsbQ were accumulated at their normal levels, and were present as unassembled units in the thylakoid lumen.

RNAi knockdown studies in *Arabidopsis* and Tobacco indicated that they could grow under low levels of PsbP (Ifuku et al., 2005b; Yi et al., 2007). When the PsbP levels are <5%, plants showed certain defects including slower photoautotrophic growth. Further analysis suggested that low levels of PsbP were sufficient to accumulate usual levels of PSII reaction center proteins,

however, formation of PSII dimers was significantly decreased. Additionally, total loss of LHCII-PSII super-complexes was also observed (Ido et al., 2009). Furthermore, reduction of PsbP by strong suppression to less than 1% led to severe defects. Since Arabidopsis has of two genes for PsbP (*psbp1 and psbp2*), interestingly enough, only PsbP1 is essential for photoautotrophic growth, and not PsbP2 (Yi et al., 2007). Complete knockdown of PsbP showed that the core proteins were highly reduced and the photoautotrophy was entirely lost, which was deduced by their growth only in presence of sucrose, and the plants are seedling lethal (Yi et. al, 2005). It was also identified that the levels of PsbP1 are reduced in mature plants. This led to further investigation on the role of PsbP. It was later identified that the PsbP is required for the early stages of seedling development, and is not required in mature plants (Allahverdiyeva et. al, 2013).

Higher Plant PsbQ in PSII

The X-ray crystal structure of soluble spinach PsbQ (PDB ID: 1VYK, (Balsera et al., 2005)) is illustrated in FIGURE 1.6. PsbQ has a four helix bundle at its C-terminal region, and the N-terminus has four prolines ^{9}P-^{12}P forming a polyproline type II or left handed helix structure. Residues ^{1}E-^{45}P form the N-terminus of PsbQ and residues ^{46}T-^{149}G form the C-terminal domain. Residues ^{14}S-^{33}Y are unresolved in this crystal structure. The FIGURE 1.10 compares the X-ray crystal structures from *T. elongatus*, *T. vulcanus* and spinach PsbQ. The structure of PsbQ has also been elucidated using NMR (PDB ID: 2MWQ, (Rathner et al., 2015)). This structure (not shown in the figure) also reveals a four helix bundle in its C-terminal domain, however, the N-terminal region is slightly different from the 1VYK crystal structure. The NMR structure has a completely resolved N-terminus. Nevertheless, the structures representing the bound state of PsbQ need further investigation. The N-terminus of PsbQ is important for its binding to PSII (Kuwabara et al., 1986). Mutant of PsbQ with its N-terminal 12 residues truncated (Δ12 PsbQ) was unable to bind to PSII

and its regions of interactions were not identified. It is also not known if PsbQ interacts with PSII just with its N-terminus or if the C-terminal domain also has a role in its interaction to PSII. Modification of a few residues in the C-terminal domain (^{90}K, ^{96}K, ^{101}K and ^{102}K), located on one face of the protein, with EDC and glycine methyl ester, did not lead to high affinity binding to PSII (Meades et al., 2005), suggesting an important role of C-terminus in the binding of PsbQ to PSII.

Although spinach contains only one copy each of the PsbP and PsbQ proteins (which interact with the PSII at stoichiometric levels), *Arabidopsis* contains two copies each of PsbP (mentioned earlier, *psbp1* and *psbp2*) and PsbQ (*psbq1* and *psbq2*). Plants lacking both *psbq1* and *psbq2* showed normal growth rates, in contrast to insertional mutants of *psbp1* (which could not grow photoautotrophically). Additionally, the triple mutants of *psbq1/psbq2/psbr* were also able to grow normally (Allahverdiyeva et al., 2013), suggesting that PsbQ (and PsbR) do not play a major role in overall growth of the Arabidopsis plant. Taking a closer look at the PSII function, complete loss of PsbQ (double mutant psbq1/psbq2) had a minor effect on the PSII function, while the triple mutant had more obvious defects in PSII activity.

The function of PsbQ in photosystem II is still unclear. The RNAi knockdown of PsbQ in Arabidopsis was studied by silencing of the genes *psbq1* or *psbq2* (Allahverdiyeva et al., 2013). These mutants didn't have any effect on the overall growth of the plant, however the single mutants and *psbq1/psbq2* double mutant had a significant effect on the levels of PsbP and PsbR CP43 in *psbq2* mutant (~85% of the wild type plant). Also, except for D1 in addition to Lhca2 (PSI) and Lhcb2 (PSII), all other PSII core proteins were similarly affected in the triple mutants. The D1 protein remained unaffected even in case of the double mutant, extrinsic subunits. Complete knockdown of PsbQ and PsbR in a triple mutant led to total loss of PsbP in a 4-week old plant. Surprisingly, none of the core proteins of PSII were affected except *psbq1*/*psbq2*. Further studies on

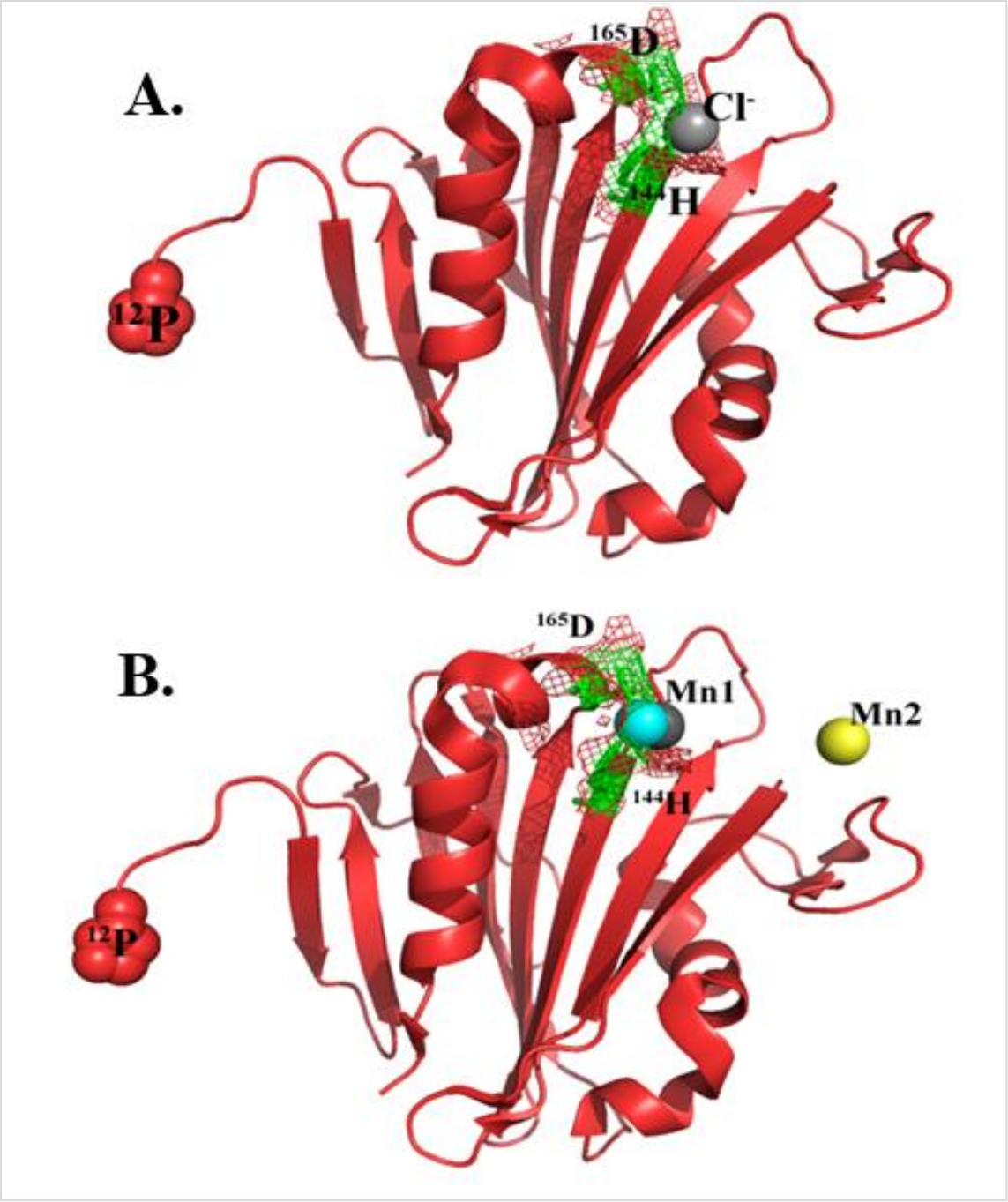


FIGURE 1.10. The X-ray Crystal Structure of PsbP from Spinach (PDB ID: 4RTI).

A. PsbP structure shows the location of residues ^{144}H, ^{165}D and Cl^- ion. **B**. Binding location of Mn^{+2} ions, Mn1 (tightly bound Mn), and Mn2 (loosely bound Mn), Mn1 is located near the residues ^{144}H and ^{165}D, and Mn2 near the residue ^{98}D (not shown in the figure).

the double mutant identified that PsbQ is required for the normal growth phenotype of plants under low light conditions (Yi et al., 2006).Additionally, loss of PsbQ (and PsbR) had significant effects on the super-complex formation. Compared to the wild type plants, the LHCII-PSII complexes were reduced by 58% and 54% in the double and triple mutants, respectively (Allahverdiyeva et al., 2013). These results emphasize the role of extrinsic proteins in super-complex formation. The TABLE 1.2 provides a brief description on the phenotype observed due to loss of different extrinsic proteins in higher plants (Bricker et al., 2013).

In vitro studies on PsbQ have usually been performed on spinach PSII membranes. The removal of PsbQ leads to defects in oxygen evolution under low Cl^- conditions (Miyao and Murata, 1985). As mentioned earlier, N-terminally truncated mutant (Δ15 PsbP) was unable to restore the oxygen evolution activity (when reconstituted onto the NaCl washed PSII membrane) under physiological levels of Ca^{+2} and Cl^- conditions (Ifuku and Sato, 2001, 2002). The oxygen evolution ability was significantly restored in the presence of PsbQ. Similar effects were observed in FTIR studies. The Amide I features of NaCl-washed PSII were not restored to normal in presence of the mutants Δ15 PsbP and H144A PsbP, these amide I bands were almost restored to normal (as in wild type) when these mutants were reconstituted in presence of PsbQ along (Kakiuchi et al., 2012). The X-ray crystal structure of soluble PsbQ (PDB ID: 1VYK, (Balsera et al., 2005)) reveals a four helix bundle at its C-terminal region, and the N-terminus has four prolines ^{9}P-^{12}P forming a polyproline type II or left handed helix structure. Residues ^{1}E-^{45}P form the N-terminus of PsbQ and residues ^{46}T-^{149}G form the C-terminal domain. Residues ^{14}S-^{33}Y are unresolved in this crystal structure. The structure of PsbQ has also been elucidated using NMR (PDB ID: 2MWQ, (Rathner et al., 2015)). This structure also describes a four helix bundle in its C-terminal domain, whereas, the N-terminal

TABLE 1.2. Effects of the Loss of Extrinsic Proteins in Different Species.

The effects of the loss of extrinsic proteins PsbO, PsbP and PsbQ along with Cyano P and CyanoQ are presented here for the organisms, Synechocystis, Chlamydomonas and Arabidopsis or tobacco.

Organism	Absent protein	Effects on PS II assembly
Synechocystis 6803	PsbO	Loss of PS II dimers
	cyanoPsbP	No effect on cyanoPsbQ accumulation
	cyanoPsbQ	No effect on cyanoPsbP accumulation
Chlamydomonas reinhardtii	PsbO	PsbP and PsbQ accumulate to near normal levels but are not assembled into the photosystem. CP47, CP43, D1 and D2 are strongly depleted apparently due to instability of the assembled photosystem
	PsbP	PsbO and PsbQ accumulate to near normal levels but are not assembled into the photosystem. CP47, CP43, D1 and D2 depleted to 30 - 65% wild - type levels
	PsbQ	No mutants available
Arabidopsis/Nicotiana	PsbO	Loss of PsbQ, and to a lesser extent, PsbP; CP43 and D1 strongly depleted, CP47 only modestly affected
	PsbP	Loss of PsbQ, PsbO slightly affected; Upon moderate suppression (5-10% wild - type levels). CP47, CP43 D1 and D2 accumulate to near normal amounts; LHCII-PS II supercomplexes do not form, LHC and core proteins not phosphorylated; Upon strong suppression (< 1% wild-type levels), CP47 and D2 do not accumulate; Defective thylakoid assembly under both moderate and strong suppression
	PsbQ	No effects under normal growth light regimes; Under low light growth regimes, CP47, CP43 and D2 strong depleted, D1 only modestly affected

Bricker et al., 2013

region is slightly different from the 1VYK crystal structure. The NMR structure has a completely resolved N-terminus.Nevertheless, the structures representing the bound state of PsbQ need further investigation. The N-terminus of PsbQ is important for its binding to PSII (Kuwabara et al., 1986). An N-terminal 12 residues truncated mutant of PsbQ (Δ12-PsbQ) was unable to bind to PSII,

suggesting its N-terminal interacts with PSII subunits. However, interaction of PsbQ to PSII requires both N-terminus and C-terminus or other domains of this protein is not known. The EDC and glycine methyl ester modification of the C-terminal domain of PsbQ (^{90}K, ^{96}K, ^{101}K and ^{102}K), located on one face of the protein, could not yield a high affinity binding to PSII (Meades et al., 2005). Thus suggesting an important role of C-terminus as well in the binding of PsbQ to PSII.

TANDEM MASS SPECTROMETRY AND CHEMICAL-CROSSLINKING IN PROTEIN CHARACTERIZATION

1.5 Introduction to Mass Spectrometry

Mass Spectrometry (MS), an analytical technique that enables structural and biophysical analysis of macromolecules present in biological samples. This method is extensively used to study plant metabolites, their site of production and quantitation, and identification of the intermediates from biosynthetic pathways. Mass Spectrometry imaging (MSI) methods are greatly used to study mechanisms under plant stress, plant defense, plant surface metabolites (produced during biotic or abiotic stress), nitrogen fixation, phosphorous deficiencies etc. MS particularly has greatly improved from the time of its genesis, about four decades ago. In addition to the above mentioned applications, MS is widely used in the "omics' field like proteomics, genomics, lipidomics, transcriptomics, glycomics etc. Protein mass spectrometry and its applications in Photosystem II are discussed below.

Mass Spectrometry when coupled to a collision or fragmentation chamber provides information on proteins at their amino-acid level. This method requires more than one mass analyzer instruments and hence is called Tandem Mass Spectrometry. Tandem MS is used to identify low-resolution three-dimensional molecular structure of a protein, any non-covalent interactions, posttranslational modifications or even amino acid sequence (peptide sequencing). The structural and functional characterization of various interactions helps in understanding the mechanics of cellular processes. The complexity in studying soluble or membrane proteins stems due to their

dynamic nature, abundance of protein expression or degradation and their variability under different environmental conditions etc. MS in such cases is very helpful to capture transient phases of proteins, proteins expressed at low levels, or to capture their dynamic states.

A schematic for the overall process of tandem mass spectrometry is illustrated in the FIGURE 1.11. The example shown in this figure has Photosystem II protein complex as the starting material, composed of several protein subunits. These proteins, when run on a one-dimensional denaturing (SDS or LiDS) poly-acrylamide gel electrophoresis (PAGE), separates proteins based on their molecular weights. The protein-band of interest is excised from the coomassie blue stained gel and subjected to enzymatic proteolysis. The resulting peptides are separated by High-performance liquid chromatography (HPLC), and the samples are directly injected into a mass spectrometer. Here the sample or peptides are ionized. Two soft ionization methods are mostly used in photosynthesis research, Electrospray Ionization (ESI) and Matrix assisted laser desorption ionization (MALDI).

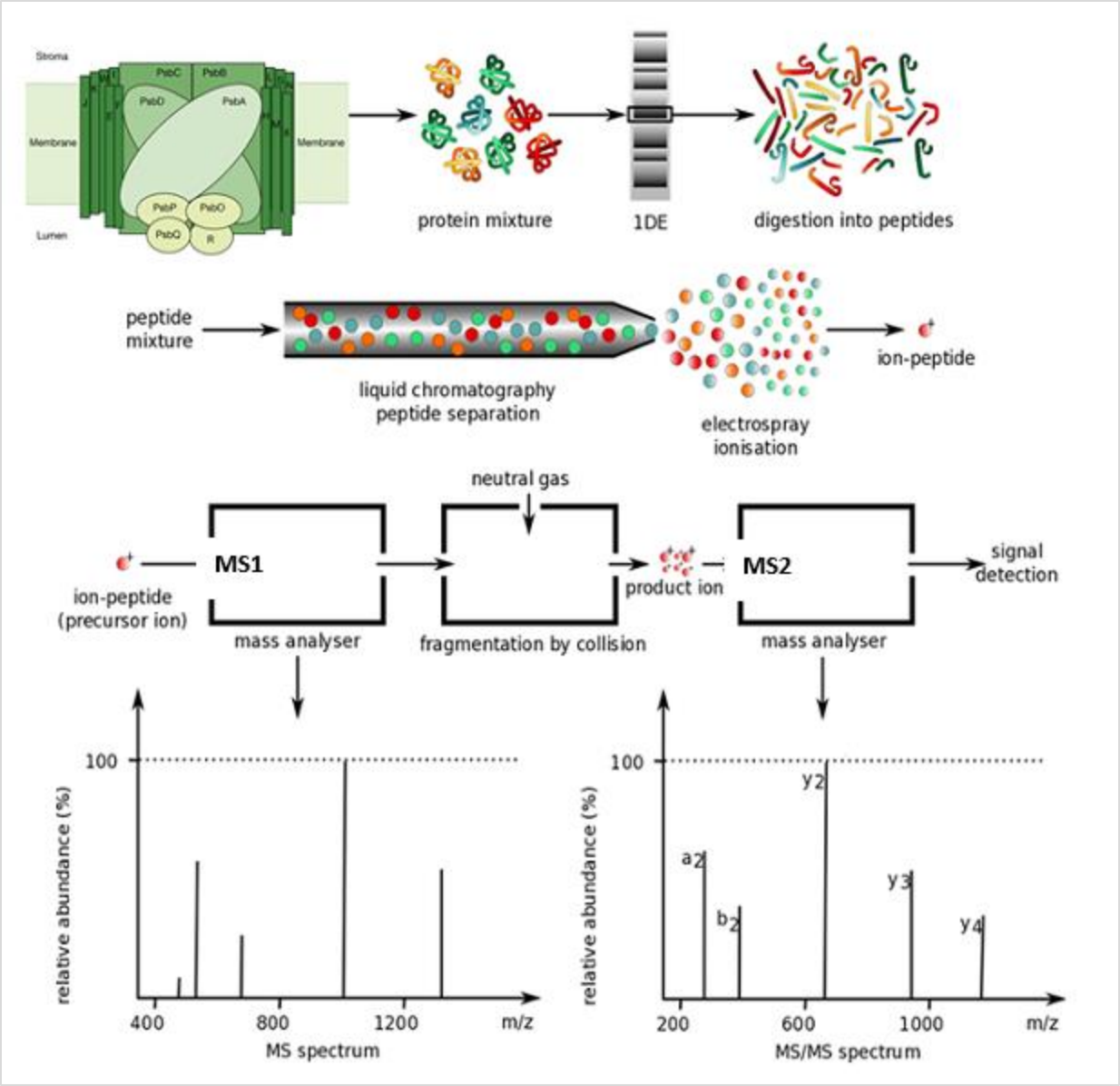
1.6 Ionization Methods

Electrospray Ionization

In (ESI)-MS, thermally labile samples solubilized in a buffer or solvent (usually polar), are introduced into the ionization source by means of a fine needle or capillary at a very low flow rate, resulting in a continuous spray of the sample. A high electrical potential applied to the needle results in the formation of an aerosol of highly charged droplets that are directed and driven by means of an electrical field while a warm gas (usually nitrogen) aids in vaporization. The presence of a sheath gas (nitrogen) around the capillary results in superior nebulization. The charged analyte enters into a sampling cone or the mass analyzer. ESI is considered to be a soft ionization technique for biomolecules, which means that the ionization process transfers minimum internal energy to the analyte.

Matrix Assisted Laser Desorption / Ionization (MALDI)

MALDI is another soft ionization technique, where samples are desorbed from a solid matrix, unlike ESI. The sample is first solubilized in a suitable solvent, to which an excess amount of the suitable matrix is added, spotted on to a MALDI plate and dried (in air or nitrogen) to solidify the matrix. The sample is thus crystallized along with the matrix. The matrix containing the sample is hit with a laser (usually nitrogen, at 337 nm), resulting in the absorption of the laser energy by the matrix followed by desorption, which in turn causes ionization of the sample into the gas phase. As in the case of other MS analyses, the ionized molecules are driven by electric and/or magnetic fields resulting in their separation based on their mass and charge and are finally detected by the detector. MALDI may be operated under vacuum or more recently, atmospheric pressure (El-Aneed et al, 2009). While both ESI and MALDI are often used in research, a choice between the two is made based on several factors as there exist distinct differences between the two. As mentioned earlier, both ESI and MALDI are soft ionization methods and are capable of detection when samples are at very low concentrations. However, the sample state in each is different-solution in ESI and a solid matrix in MALDI. This is advantageous when a liquid chromatograph (LC) is attached to the MS and the sample has to be introduced in liquid state. While ESI is known to yield reproducible data, it should be noted that ions are multiplied and hence the absorbance in the chromatogram does not reflect the true sample concentration.



Modified from Pharmchem.ucsf.edu

FIGURE 1.11. Overview of Tandem Mass Spectrometry.

Steps involved in tandem mass spectrometry process are shown here. MS1 refers to first Mass Spectrometer and MS2 is the second mass spectrometer. Spectra are illustrated as m/z on the X-axis with relative abundance on the Y-axis.

MALDI is a fairly robust method and allows for presence of salts/detergents at moderate levels as the ionized samples can escape the contaminants similar to how they escape the matrix molecules. However, in ESI, these can be contaminants and can compete with the sample for ionization, affecting the results.

The ionized samples travel through the chamber based on their m/z ratios and are analyzed by the MS1 analyzer (See FIGURE 1.11). Of all the mass analyzers (linear quadrupole ion trap, three-dimensional quadrupole ion trap, magnetic-, electric- and sector mass analyzers, time-of-flight mass analyzer, ion cyclotron resonance (ICR) that uses electric/magnetic fields), ESI is generally attached to ion trap mass analyzers, more specifically a linear ion trap (Banerjee and Mazumdar, 2012). Ion traps can hold ions for a longer periods, assisting in fragmentation of the parent ions MH^+. Although, it can determine the molecular mass accurately, obtaining site-specific information on structure, modifications of certain residues etc. from its fragment ions is not possible. Hence, MS/MS (MS2) is employed to identify such details.

1.7 Precursor Ion Fragmentation Methods

For MS2, the parent ions (single or multiply charged) from electrospray are further fragmented in the mass analyzer, the linear ion trap in this case also contains helium gas, when a small radiofrequency is applied, these ions thus move faster, experiencing resistance with the helium gas (He), and sample ions colliding with an inert gas causes fragmentation of the sample ions, thus converting sample ion's kinetic energy into internal energy (Shukla and Futrell, 2000; Sleno and Volmer, 2004). This process is known as Collision-induced dissociation (CID) or collisionally activated dissociation (CAD). CID can be performed at low collision energies (eV range) or high collision energies (k eV range). For low collision energies, tandem quadrupoles or ion trap methods are employed, while for higher collision energy, time-of-flight (TOF) instrument is used (Sleno and

Volmer, 2004). As the name suggests, TOF measures m/z of ions based on the time taken for these ions to reach the detector at a given distance. These ions are accelerated in an electric field, which results in the ions with the same charges to have similar kinetic energy. Hence, these ions reach the detector based on their m/z ratios. For example, smaller peptides can reach much faster than the larger peptides when they have similar charges. One advantage of a TOF analyzer is that, all the ions eventually reach the detector and hence it can detect ions a large range of masses. Thus, TOF is generally connected to ESI or MALDI for detection of macromolecules (>mega Dalton mass range). Many advanced TOF analyzers have a resolving power >12,000 and mass accuracy of low ppm are achievable (El-Aneed et al., 2009). Fourier Transform (FT) ICR uses magnetic and electrical fields (Orbitrap) to generate ions, transport them to the trap and help in excitation of ion motion along with their detection in the electrodes of the ion trap. An ion experiences a centripetal force on entering the magnetic field, and hence it moves in a circular motion in an orbit. The frequency of an ion's orbit is detected and shown as spectra of m/z ratios vs abundance of the ions.

Mass accuracy is a very important parameter, especially when dealing with fragmented ions. Higher mass accuracies are obtained with high-resolution mass analyzers like Fourier transform ion cyclotron resonance (FTICR) and orbitrap. These are often coupled with ion trap mass analyzers in MS2 or MS3 mass spectrometric analysis. Range of mass accuracies obtained with different mass analyzers is illustrated in the TABLE 1.3. Resolution is the other feature most required during analyses of the samples. A high resolution is also required to near accurately measure the mass and quantify the abundance, which would not be possible if the two peaks from different constituents are not sufficiently resolved (Scigelova et al., 2011). Mass tolerance of a mass analyzer corresponds to its resolution. The resolution is generally affected when different constituents or peptides from a sample mixture have similar peaks (similar mass to charge ratios), and thus, if resolution is not high

enough, these constituents are categorized in the same peak, and thus inefficient analysis of the sample results. Hence, higher mass accuracies are required while studying samples with multiple

TABLE 1.3. Mass Accuracies and Sensitivity of Mass Spectrometers used in Proteomics.

These values for mass accuracies are found under optimal conditions. Practically however, observed mass accuracies are typically somewhat lower. FTICR, Fourier-transform ion cyclotron resonance, LTQ-Orbitrap; linear trap quadrapole; TOF, time-of-flight; TQ, triple quadrapole; and LIT, linear ion trap.

Instrument type	Mass accuracy	Sensitivity
FTICR	0.1 - 1 ppm	Femtomole
Orbitrap	0.1 - 1 ppm	Femtomole
TOF	1 - 5 ppm	Femtomole
TQ	50 - 200 ppm	Femtomole
LIT	50 - 200 ppm	Femtomole

Bricker et al.. 2015

proteins, protein complexes, post-translational modifications or protein crosslinking. The fragmentation generates b- and y-type ions, with different charge states. The other most used fragmentation method is electron capture dissociation (ECD), used to study non-covalent interactions and can be easily coupled to FTICR mass analyzer. ECD produces c- and z-type ions. The FIGURE 1.12 shows the fragmentation and the ions generated in the process. One mass analyzer analyzes the peptides of a protein, and the second mass analyzer further analyzes the fragmented ions. The most common type of ions produced from ion traps and q-TOF mass spectrometers are a-, b- and y-type produced from low energy collision. The b-type ion series extend from the N-terminus while the y-type series from the C-terminus. Some fragmentations occur more readily than the others, and hence, these ions are not generated for each peptide bond. The parent

peak can be predicted for its b- and y-type ions, and mass analysis for the MS/MS analysis can be done manually, although, automation makes it much faster.

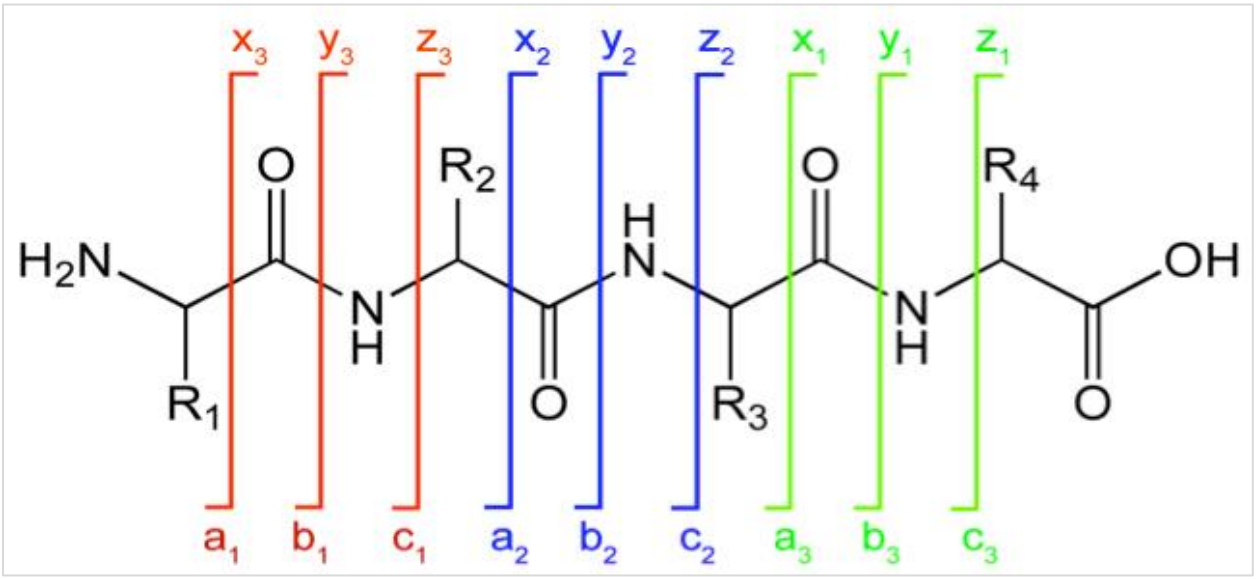


FIGURE 1.12. Ion Generated upon Fragmentation with CID.

An example peptide labelled with different ion types is shown in the figure. b- and y- ion types are generated by CID by the cleavage of the amide bond along the peptide backbone are preferentially important for sequence determination. Other ions a-, c-, x-, and z- types are also shown in the figure. The ions arising from the N-terminus are b-ion series while the ions generated from the C-terminus are termed y-ion series.

1.8 Chemical Crosslinking of Proteins

The knowledge of interaction in protein-complex assemblies is critical to understanding the underlining biological processes. Chemical protein-crosslinkers function by linking the functional groups on the protein with the reactive groups on the crosslinker. These crosslinkers form a covalent bond/peptide bond between the crosslinker and the functional groups, thus linking the two functional groups by its carbon chain or in case of a zero-length crosslinker; an amide bond is formed between the two functional groups. Thus the crosslinkers add distance constraints to the structure of the protein, and identification of these crosslinked-residues helps in determining the interacting domains

and three-dimensional structure of the proteins. Protein-crosslinking coupled to tandem mass spectrometry is an excellent tool to provide low-resolution models of proteins and protein complexes. Due to certain drawbacks to X-ray crystallography and NMR methods mentioned earlier, crosslinking procedures may be very useful. The chemical crosslinkers used in the PSII research are NHS-biotin, glutaraldehyde, disuccinimidyl suberate (DSS), Dithiobis (succinimidyl propionate) (DSP), BS3 and 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide • HCl (EDC).

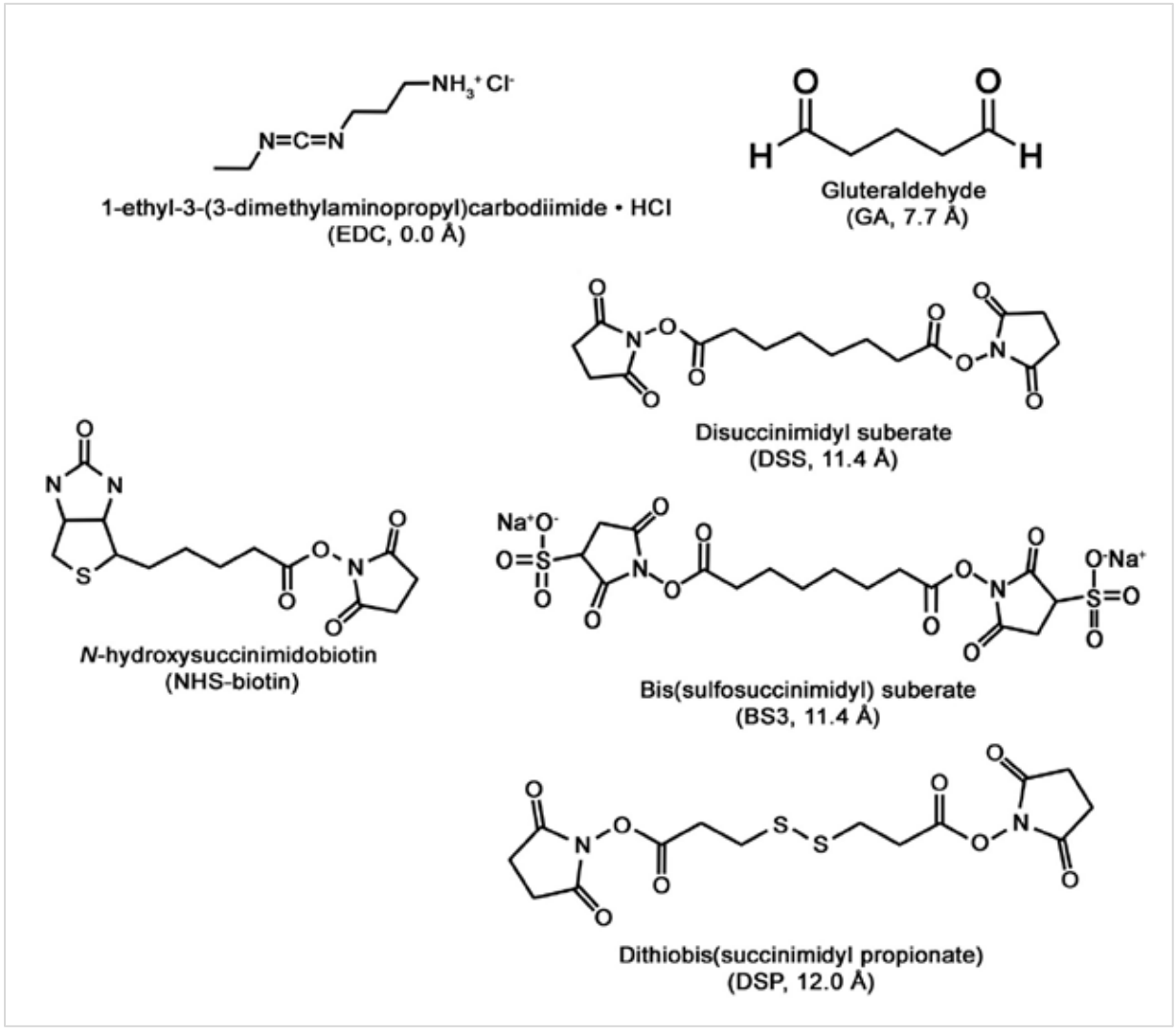


FIGURE 1.13. The Structures of Crosslinkers Used in PSII Studies.

Crosslinkers with their chemical names and structures are illustrated, their common names are shown in parentheses along with the distance between the residues they can crosslink.

Structures of these protein-crosslinkers are illustrated in FIGURE 1.13. These crosslinkers modify certain surface residues, and the mass changes with their labelling can be determined by tandem mass spectrometry. Gluteraldehyde is a five-carbon compound that reacts with primary amines readily at neutral pH, although it can also react with other groups like imidazole, thiol and phenols at other pH conditions. Its reactivity is dependent on nucleophiles and hence it can react with tyrosine, tryptophan, phenylalanine, histidine, cysteine, proline, serine, glycine, arginine residues (Migneault et al., 2004). Gluteraldehyde changes structurally depending on the pH, its most reactive form at a given pH is unknown, and hence the reaction mechanism is not known too. Another crosslinker, formaldehyde, displays reactivity with these amino acids similar to glutaraldehyde, and is an emerging crosslinking reagent to study protein-interactions these days (Klockenbusch et al., 2012; Sutherland et al., 2008; Toews et al., 2008).

The NHS-crosslinkers are the most commonly used homobifunctional chemical modifiers. They readily react with primary amines (ε amino group of lysine, N-terminal amine of the protein) at a basic pH (pH 7.0 - 9.0), and at acidic pH (pH 6.0), it can react with tyrosyl-, threonyl-, and seryl-residues (Kalkhof and Sinz, 2008a; Madler et al., 2009). Three of the crosslinkers with NHS-head groups are DSS, DSP and BS3. Their functional head groups are separated by an eight-carbon chain and can crosslink residues up to 11.4 Å apart. During reaction, the crosslinker forms a stable amide bond between the residues with the removal of the functional head groups. DSS and DSP are soluble in organic solvents, and are generally dissolved in DMSO prior to adding it to the aqueous reaction mixture. BS3 on the other hand has an additional sulfo group, which makes it water soluble, and cannot cross the biological membranes, unlike DSS and DSP. Thus BS3 is not suitable for *in vivo* crosslinking experiments.

EDC is a zero-length carbodiimide that is water soluble and can activate carboxyl residues for their reaction with primary amines, forming O-acylisourea as a reaction intermediate, to form an amide bond with the release of isourea as a by-product. EDC can crosslink the carboxy- group of glutamate and aspartate to the side-chain primary amine of lysine and the amino group at the N-terminus. The crosslinker can readily react in a pH range of 4.5 – 7.2, and can only crosslink the residues within their van der Waals distance.

In addition to protein-crosslinking, EDC and NHS-biotin, were used earlier as the surface residues modifying agents. The protein complexes were treated with these reagents and surface modifications were identified by tandem mass spectrometry. Some domains of the protein are not accessible to these chemical reagents possibly because theese are either hidden or buried, allowing us to identify the surface residues, buried regions and/or interacting domains in a protein complex. A much reliable technique used these days is radiolytic footprinting. This procedure can globally oxidize residues that are surface exposed in a protein complex. The method involves production of hydroxyl-radicles upon exposure to X-rays. The hydroxyl radicles are generated from water present in the sample buffer. These free radicles are highly reactive and can modify several residues exposed on the surface of the proteins, accessible to the bulk solvent. Amino acids in the relative order of their reactiveness to oxidative modification have been described earlier (Takamoto and Chance, 2006).

1.9 Tandem Mass Spectrometry in PSII Research

The site of water oxidation in photosystem II is made of Mn_4CaO_5 cluster along with the ligands offered by intrinsic and extrinsic membrane proteins. The extrinsic membrane proteins (PsbO, PsbP and PsbQ) help in stabilizing this Mn_4CaO_5 cluster and protecting it from the exogenous reductants.

The D1 protein in PSII undergoes constant damage-repair process. PSII is most susceptible to light induced damage, and the protein D1 is principally prone to ROS damage (Pospisil, 2009; Pospísil et al., 2004). It is hypothesized that during water oxidation process, and release of oxygen, several ROS radicles ($O2^{-}$, $^{\bullet}OH$, H_2O_2) are generated. Damage metal clusters may also lead to more ROS production. These sites of ROS production have been studied by tandem mass spectrometry. Several oxidized residues on D1 and D2 proteins were oxidized near $Pheo_{D1}$ and Q_A sites, and other residues on CP43 near Mn_4CaO_5 cluster, suggesting that they are the source of ROS generation (Frankel et al., 2013; Sharma et al., 1997).

Additionally, water has to travel to the Mn_4CaO_5 cluster for its oxidation, hence, several water / oxygen channels were proposed computationally based on the cyanobacterial crystal structure (Ho, 2012; Ho and Styring, 2008; Murray and Barber, 2007; Petrek et al., 2006; Vassiliev et al., 2010) These studies were experimentally tested by exposing the PSII membranes to X-rays for 0, 4, 8 and 16 seconds followed by tandem mass spectrometry. Numerous oxidized residues were identified in this time course experiment (Frankel et al., 2012). The modified residues were found in D1, D2, CP47 and CP43. Some of these residues were only oxidized when irradiated with X-rays, and these residues were similar to the computationally identified water channel. These residues seemed to form a continuous track of oxidized residues channeling from the surface of the PSII complex to the metal cluster. The residues, which were modified in the absence of irradiation, were referred to as ROS/O2 exit channels.

Radiolytic footprinting followed by tandem mass spectrometry can identify oxidatively modified residues by exposing the PSII proteins to radiation (X-rays). This method can modify the surface residues exposed to the bulk solvent (Mummadisetti et al., 2014). The regions not modified or oxidized, are those which are buried within the protein complexes and are indicative of their

interaction with other proteins in this region. Furthermore, phosphoproteomic studies help in identifying the post-translational modifications in a protein. Several PSII proteins including PsbP-1, PsbP-2 and PsbQ-2 (in *Arabidopsis*) were phosphorylated based on the studies from Arabidopsis and Chlamydomonas (Hansson and Vener, 2003; Puthiyaveetil and Kirchhoff, 2013; Reiland et al., 2009; Tikkanen and Aro, 2012; Turkina et al., 2006; Vener, 2007; Vener et al., 2001).

Protein crosslinking methods in PSII research have been in use for more than 30 years. However, they have been coupled to high-resolution tandem mass spectrometry much recently to identify the crosslinking sites. This method has been useful in identifying the location of certain subunits not present in the X-ray crystal structure, and hence it can provide low-resolution models on the subunit composition and their location in a protein complex. Only the examples from the extrinsic proteins of PSII are included here.

Cyanobacterial protein CyanoQ (structural homolog of PsbQ from higher plants) is a subunit of PSII. This protein was identified as a part of the PSII complex, when a histidine tag at the C-terminus of CyanoQ could extract the entire PSII complex (see section 1.4.1 for details) by affinity chromatography (Roose et al., 2007a). Protein crosslinkers EDC, BS3 and DTSSP when used to modify cyanobacterial PSII, extracted from the C-terminally-His-tagged CyanoQ, identified several inter- and intra- molecular crosslinked residues (Liu et al., 2014). The inter-molecular crosslinked residues were identified between CyanoQ::PsbO (CyanoQ:^{120}K - PsbO:^{180}K and CyanoQ:^{120}K - PsbO:^{59}K), CyanoQ::PsbB (CyanoQ:^{102}K - PsbB:^{440}D). Several intra- molecular crosslinked residues from CyanoQ, identified in this study, were completely consistent with the X-ray crystal structure. However, one of the crosslinked residue pair CyanoQ:^{96}K - CyanoQ:^{120}K spanned 41 Å (with the crosslinker BS3) on the CyanoQ crystal structure. It was hypothesized that these crosslinking pairs were interacting from two different CyanoQ molecules. Hence, these dimeric

CyanoQ were placed between two PSII monomers in a symmetrical antiparallel fashion associating from the lumenal face of PSII. Additionally, each CyanoQ was placed between PsbO and PsbB within a PSII monomer. Similar examples in identifying the interacting regions of PSII subunits were seen for the higher plant PsbP from green alga and spinach. The EDC crosslinker could crosslink PsbP and PsbE in the green alga Chlamydomonas (Nagao et al., 2010) and the crosslinked residues among these proteins were identified in spinach using tandem mass spectrometry (Ido et al., 2014). In this experiment, the residue pairs PsbP:^{1}A – PsbE:^{57}E were identified in multiple independent spectra. The authors also identified crosslinked residues between PsbP::PsbR (PsbP:^{27}K- PsbR:^{22}D), a tertiary peptide containing two PsbP peptides and one peptide from CP26, PsbP::CP26::PsbP (PsbP:^{115}E-PsbP:^{173}K-PsbP:^{174}K-CP 26:^{96}E). These results suggested that PsbP closely interacts with PsbR, PsbE and CP26.

1.10 Analysis of Crosslinked Residues from Mass Spectrometry Data

When studying a mixture of proteins, the number of peptides generated from the sample, the ions produced after their fragmentation and consequently the m/z ratio of each of those ions generated are several thousands in number which are extremely strenuous for manual sorting. More complexity is added to this when not just b- and y- type ions but also the ions from the side chains are generated (even worse, if the backbone is cleaved multiple times). Thus, the analysis of site-specific sequencing using tandem mass spectrometry is always automated by a statistical software.

Although, de novo MS/MS analysis of the peptide sequence can be done from a mixture of proteins, a more reliable and reproducible result can be generated if a protein library is provided, and the MS/MS data are searched against this protein library. In all the currently available softwares, the user provides a library of target proteins from which the theoretical peptides and subsequently theoretical spectra are generated for a specific enzyme. The data from tandem mass spectrometry

are compared to these theoretical spectra to identify the location of chemical or post-translational modifications. There are several softwares available to analyze the mass spectrometry data. A few of these include MASCOT (Perkins et al., 1999), GPMAW (Peri et al., 2001), SEQUEST (Eng et al., 1994) and MassMatrix (Xu and Freitas, 2007b, 2009) for post-translational modifications, pLink (Yang et al., 2012), xLink (Kosinski et al., 2015), MassMatrix and StavroX (Gotze et al., 2012) for chemical or oxidative modifications and protMapMS (Kaur et al., 2009) specifically for oxidative modifications. For my dissertation, MassMatrix and StavroX have been used for the mass analysis of BS3 and EDC modified proteins or peptides and are described below.

StavroX

StavroX is a fairly new tool developed by Michael Gotze (Gotze et al., 2012) that has not been explored much in photosynthesis research. This search engine uses a non-probabilistic scoring method for crosslinked proteins. It is available free of cost, and runs without installation on windows platforms. StavroX uses a FASTA library to generate theoretical peptide mass list for a specific enzyme, which is used to compare with the data generated from MS/MS. Enzymatic cleavage sites are stated with a question mark (?) on either left- or right- hand side of the single letter code of the amino acid depending on the N- or C-terminal cleavage by the protease, for example, Trypsin digests on the C- terminus of lysine or arginine, and hence is represented as K?, R?, and since proline prevents the activity of trypsin at that site, letter P for proline is stated in 'blocked by' field. Lower case letters indicates oxidative modifications, e.g. modification of methionine (M) would be (m) for oxidized methionine.

Currently the only acceptable file-format for Mass Spectrometry input data files for StavroX is Mascot generic file (.mgf). Precursor and product ion mass measurements are defined based on the type of MS/MS method used, just like in Mass Matrix. The cross linkers used for the experiment

are defined by their elemental composition on StavroX. A decoy search library, by reversing the sequences of the protein library provided, is used in each case to identify the false discovery rate for a particular search. Based on several searches, it was identified that the scores larger than 100 have false positive rate less than 2%. One greatest advantage of StavroX over MassMatrix is that it can have any number of proteins in its library for a single search. MassMatrix webserver had a similar feature, which is unavailable at present, and the PC version can only search a maximum of three proteins to identify crosslinks in a single search. It is easier to sort data with StavroX software first, to identify the proteins in the reaction mixture, followed by the using appropriate search parameters in MassMatrix to reduce its number of searches. The scoring method of StavroX is much different from MassMatrix and hence the crosslinks obtained cannot be compared based on their scores. A score in StavroX at a given FDR represents the chance of occurrence is 95.5%, for example, if the FDR for a crosslink is 20, the chance of occurrence of that crosslink would be 95.5%. However for our crosslinking data we used a much stringent criteria of 3X FDR, so from previous example, a score of 60 (3x 20 = 60) is considered, also the other criteria set was to have a score above 100, to reduce the probability of false discovery. So, in the same example, this crosslink would be considered only if the score of the crosslink is >100.

StavroX is considered one of the good programs in the market today, to identify the maximum number of crosslinked peptides in a single search and shortest time span, however, it has a few limitations as well. One of the drawbacks is that this program has not been optimized well to relate its scores to the probability of occurrence of a crosslink to get its *P* values. Additionally, it reports crosslinks of lysine with the C-terminal end of a tryptic peptide, since trypsin cannot cleave at a lysine that is crosslinked, and hence the crosslinks have to be carefully chosen while using this software (Sriswasdi et al., 2014). Its FDR increases with an increase in the number of proteins in the

FASTA library, and hence it is advisable to use those proteins which were identified to participate in the formation of a crosslink product experimentally, and are potential participants in crosslink formation.

MassMatrix

It is a much-used software developed on ANSI C++ and runs on Windows and Linux machines. It supports .mzXML, .mgf, .mzDATA input file formats for tandem mass spec data. The Protein libraries have to be provided in FASTA or .BAS formats. Modifications during the search can be specified as variable and fixed modification (similar to StavroX). MassMatrix uses a probabilistic scoring method to rank peptide and protein matches (Xu and Freitas, 2007b). It considers mass accuracy while scoring the peptides, which improves the score for a peptide and successively true positives during peptide assignments in protein identification. A comparison of higher mass accuracies, suggested that MassMatrix has a better scoring system among Mascot, SEQUEST, X!Tandem and Open Mass Spectrometry Search Algorithm (OMSSA) (Geer et al., 2004; Xu and Freitas, 2009). Scoring in this program is in terms of *P* value, represented as pp, pp2 and pptag. The scores pp and pp2 refer to the negative logarithm of the probability of random occurrence of a peptide, individually however, a pp score represents the peptide match from the product ions in the experimental spectrum with the theoretical spectrum, and pp2 score refers to the abundance of these matched ions (Xu and Freitas, 2008, 2009). And the pptag score, refers to "sequence tags" or the number of consecutive 'b' and 'y' ions identified from the peptide sequence. The higher the sequence tags identified, the higher is the probability that the match is correct, and consequently and higher probability score or pp-tag value. These values are influenced by the size of the protein database used for mass spec data searches, just like in StavroX, and the mass tolerance. When more proteins are present in the protein search database (hence more proteins in the decoy

library) there are more theoretical peptides to be compared with, and statistically, there is a greater probability for incorrect peptide identifications, and with higher the mass tolerance, higher is the chance of incorrect peptide count.

The major challenge with mass spectrometry is the huge amount of data generated (as m/z values of different ions), thus it is important to use an appropriate statistical tool, that can carefully include and analyze all the data, with maximum identifications (of modified peptides), and with the highest accuracy. Thus we used two methods, one probabilistic (MassMatrix) and the other is non-probabilistic (StavroX) scoring systems, and applied much stringent criteria for data analysis.

1.11 Protein Modeling

In this dissertation, PsbO and PsbP are modelled using Homology Modelling. The models were optimized by molecular dynamics, and evaluated for their stereochemistry using Ramachandran analysis and Z-scores.

Homology Modeling (HM)

HM, also known as comparative modeling, is the method of choice to understand the three-dimensional structure of a protein based on its amino acid sequence. X-ray and NMR are very powerful techniques to obtain high-resolution structures of proteins, however these methods aren't always successful (Acharya and Lloyd, 2005; Yee et al., 2005). In the absence of experimentally determined structures, low-resolution structural information can be obtained by computational methods. Computational methods become more accurate by homology modeling, which can help in understanding the spatial arrangement of the important residues in a protein.

Evolutionarily related proteins share similar structures. HM concept was designed on a hypothesis that the structural information of a protein is generally more conserved evolutionarily than the sequence. This method of modelling based on homology involves the following steps: (1)

Search of a template based on fold recognition and sequence alignment (2) Model building based on the 3D structure of the template (3) Auto-modeling the gap regions in the template (4) Model refinement using optimization protocols (5) and Model validation

The first step in HM is identification of the most appropriate template. A protein with 50% or more sequence similarity can give rise to very similar structures (Pearson, 2013; Vyas et al., 2012). To identify the right template, sequence similarity search can be performed on a search engine like BLAST (Basic Local Alignment Search Tool), which is a very widely used and reliable way to sequence similarity assessments. In this step, an unknown protein sequence is compared to a known protein sequence whose structural information is available. Once we have the template, its primary structure's sequence is aligned to the target sequence using an alignment program like PSI-BLAST (Position Specific Iterated – BLAST) (Altschul et al. 1990).

The second step is to build model based on the fold or sequence alignment with the template. This can be done in one of the following ways: rigid-body assembly, segment matching, spatial restraint, or artificial evolution (Vyas et al., 2012). In my research we are using the spatial restraint method, which involves aligning and satisfying the spatial restraints of the target protein by defining the constraints on the target sequence. For this process, we have used the program MODELLER for comparative protein structure modelling and adding constraints based on the data obtained from protein crosslinking experiments. From on the constraints stated, this program generates several energetically favorable models. These models generate gaps or deletions based on the alignment with the template, which are generally filled with loops, however it can be changed to a desired secondary structure using similar programs.

The models generated have to be energy minimized further for their structural optimization which is done using molecular mechanics force fields followed by further refinement using one of

the following methods: molecular dynamics, Monte Carlo methods or genetic- algorithm based sampling. These models have to be validated to obtain near native structures for more reliable results which forms the basis for our fourth step. The optimized structures are analyzed for their quality by their Ramachandran plots. Only a few distortions are allowed in a good model, and thus further refinement may be required for a model to qualify. A variety of programs are available these days, of which WHAT_CHECK (Hooft et al., 1996) can be used for crystallography related problems. All models are generally validated for their stereochemistry (which includes symmetry, geometry and structural packing quality) and fitness of sequence to structure which can be studied using softwares like PROSA II (Laskoswki et al., 1993).

References

Acharya, K.R., and Lloyd, M.D. (2005). The advantages and limitations of protein crystal structures. Trends Pharmacol Sci *26*, 10-14.

Ago, H., Adachi, H., Umena, Y., Tashiro, T., Kawakami, K., Kamiya, N., Tian, L., Han, G., Kuang, T., Liu, Z., *et al.* (2016). Novel features of eukaryotic photosystem II revealed by its crystal structure analysis from a red alga. The Journal of Biological Chemistry.

Allahverdiyeva, Y., Suorsa, M., Rossi, F., Pavesi, A., Kater, M.M., Antonacci, A., Tadini, L., Pribil, M., Schneider, A., Wanner, G., *et al.* (2013). Arabidopsis plants lacking PsbQ and PsbR subunits of the oxygen-evolving complex show altered PSII super-complex organization and short-term adaptive mechanisms. The Plant Journal : for cell and molecular biology *75*, 671-684.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990). Basic local alignment search tool. Journal of Molecular Biology *215*, 403-410.

Aoi, M., Kashino, Y., and Ifuku, K. (2014). Function and association of CyanoP in photosystem II of Synechocystis sp PCC 6803. Res Chem Intermediates *40*, 3209-3217.

Balsera, M., Arellano, J.B., Revuelta, J.L., de las Rivas, J., and Hermoso, J.A. (2005). The 1.49 A resolution crystal structure of PsbQ from photosystem II of Spinacia oleracea reveals a PPII structure in the N-terminal region. Journal of Molecular Biology *350*, 1051-1060.

Banerjee, S., and Mazumdar, S. (2012). Electrospray ionization mass spectrometry: a technique to access the information beyond the molecular weight of the analyte. International Journal of Analytical Chemistry *2012*, 282574.

Barth, A. (2007). Infrared spectroscopy of proteins. Biochimica et Biophysica Acta *1767*, 1073-1101.

Bondar, A.N., and Dau, H. (2012). Extended protein/water H-bond networks in photosynthetic water oxidation. Biochimica et Biophysica Acta *1817*, 1177-1190.

Bricker, T., and Burnap, R. (2005). The Extrinsic Proteins of Photosystem II. In Photosystem II, T. Wydrzynski, K. Satoh, and J. Freeman, eds. (Springer Netherlands), pp. 95-120.

Bricker, T.M. (1992). Oxygen evolution in the absence of the 33-kilodalton manganese-stabilizing protein. Biochemistry *31*, 4623-4628.

Bricker, T.M., and Frankel, L.K. (2011). Auxiliary functions of the PsbO, PsbP and PsbQ proteins of higher plant Photosystem II: a critical analysis. Journal of Photochemistry and Photobiology. B, Biology *104*, 165-178.

Bricker, T.M., Mummadisetti, M.P., and Frankel, L.K. (2015). Recent advances in the use of mass spectrometry to examine structure/function relationships in photosystem II. Journal of Photochemistry and Photobiology. B, Biology *152*, 227-246.

Bricker, T.M., Roose, J.L., Fagerlund, R.D., Frankel, L.K., and Eaton-Rye, J.J. (2012). The extrinsic proteins of Photosystem II. Biochimica et Biophysica Acta *1817*, 121-142.

Bricker, T.M., Roose, J.L., Zhang, P., and Frankel, L.K. (2013). The PsbP family of proteins. Photosynthesis Research *116*, 235-250.

Britt, R.D., Campbell, K.A., Peloquin, J.M., Gilchrist, M.L., Aznar, C.P., Dicus, M.M., Robblee, J., and Messinger, J. (2004). Recent pulsed EPR studies of the photosystem II oxygen-evolving complex: implications as to water oxidation mechanisms. Biochimica et Biophysica Acta *1655*, 158-171.

Brzezowski, P., Wilson, K.E., and Gray, G.R. (2012). The PSBP2 protein of Chlamydomonas reinhardtii is required for singlet oxygen-dependent signaling. Planta *236*, 1289-1303.

Burnap, R.L., Shen, J.R., Jursinic, P.A., Inoue, Y., and Sherman, L.A. (1992). Oxygen yield and thermoluminescence characteristics of a cyanobacterium lacking the manganese-stabilizing protein of photosystem II. Biochemistry *31*, 7404-7410.

Burnap, R.L., and Sherman, L.A. (1991). Deletion mutagenesis in Synechocystis sp. PCC6803 indicates that the Mn-stabilizing protein of photosystem II is not essential for O2 evolution. Biochemistry *30*, 440-446.

Cao, P., Xie, Y., Li, M., Pan, X., Zhang, H., Zhao, X., Su, X., Cheng, T., and Chang, W. (2015). Crystal structure analysis of extrinsic PsbP protein of photosystem II reveals a manganese-induced conformational change. Molecular Plant 8, 664-666.

Cardona, T., Sedoud, A., Cox, N., and Rutherford, A.W. (2012). Charge separation in photosystem II: a comparative and evolutionary overview. Biochimica et Biophysica Acta *1817*, 26-43.

Cormann, K.U., Bartsch, M., Rogner, M., and Nowaczyk, M.M. (2014). Localization of the CyanoP binding site on photosystem II by surface plasmon resonance spectroscopy. Frontiers in Plant Science *5*, 595.

Darnell, S.J., LeGault, L., and Mitchell, J.C. (2008). KFC Server: interactive forecasting of protein interaction hot spots. Nucleic Acids Research *36*, W265-269.

Dau, H., Grundmeier, A., Loja, P., and Haumann, M. (2008). On the structure of the manganese complex of photosystem II: extended-range EXAFS data and specific atomic-resolution models for four S-states. Philosophical transactions of the Royal Society of London. Series B, Biological Sciences *363*, 1237-1243; discussion 1243-1234.

de Vitry, C., Olive, J., Drapier, D., Recouvreur, M., and Wollman, F.A. (1989). Posttranslational events leading to the assembly of photosystem II protein complex: a study using photosynthesis mutants from Chlamydomonas reinhardtii. J Cell Biol. *109*, 991-1006.

Debus, R.J. (2008). Protein Ligation of the Photosynthetic Oxygen-Evolving Center. Coordination Chemistry Reviews *252*, 244-258.

Debus, R.J., Barry, B.A., Babcock, G.T., and McIntosh, L. (1988). Site-directed mutagenesis identifies a tyrosine radical involved in the photosynthetic oxygen-evolving system. Proceedings of the National Academy of Sciences of the United States of America *85*, 427-430.

El-Aneed, A., Cohen, A., and Banoub, J. (2009). Mass Spectrometry, Review of the Basics: Electrospray, MALDI, and Commonly Used Mass Analyzers. Appl Spectrosc Rev *44*, 210-230.

Eng, J.K., McCormack, A.L., and Yates, J.R. (1994). An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. Journal of the American Society for Mass Spectrometry *5*, 976-989.

Ferreira, K.N., Iverson, T.M., Maghlaoui, K., Barber, J., and Iwata, S. (2004). Architecture of the photosynthetic oxygen-evolving center. Science *303*, 1831-1838.

Forbush, B., Kok, B., McGloin, M. (1971). Cooperation of charges in photosynthetic O2 evolution-II. Damping of flash yield oscillation, deactivation. Photochemistry and Photobiology *14*, 307-321.

Frankel, L.K., Sallans, L., Limbach, P.A., and Bricker, T.M. (2012). Identification of oxidized amino acid residues in the vicinity of the Mn(4)CaO(5) cluster of Photosystem II: implications for the identification of oxygen channels within the Photosystem. Biochemistry *51*, 6371-6377.

Frankel, L.K., Sallans, L., Limbach, P.A., and Bricker, T.M. (2013). Oxidized amino acid residues in the vicinity of Q(A) and Pheo(D1) of the photosystem II reaction center: putative generation sites of reducing-side reactive oxygen species. PloS One *8*, e58042.

Gabdulkhakov, A., Guskov, A., Broser, M., Kern, J., Muh, F., Saenger, W., and Zouni, A. (2009). Probing the accessibility of the Mn(4)Ca cluster in photosystem II: channels calculation, noble gas derivatization, and cocrystallization with DMSO. Structure *17*, 1223-1234.

Gallivan, J.P., and Dougherty, D.A. (1999). Cation-pi interactions in structural biology. Proceedings of the National Academy of Sciences of the United States of America *96*, 9459-9464.

Geer, L.Y., Markey, S.P., Kowalak, J.A., Wagner, L., Xu, M., Maynard, D.M., Yang, X., Shi, W., and Bryant, S.H. (2004). Open mass spectrometry search algorithm. Journal of Proteome Research *3*, 958-964.

Ghanotakis, D.F., Babcock, G.T., and Yocum, C.F. (1984). Calcium reconstitutes high rates of oxygen evolution in polypeptide depleted Photosystem II preparations. FEBS Letters *167*, 127-130.

Glockner, C., Kern, J., Broser, M., Zouni, A., Yachandra, V., and Yano, J. (2013). Structural changes of the oxygen-evolving complex in photosystem II during the catalytic cycle. The Journal of Biological Chemistry *288*, 22607-22620.

Gotze, M., Pettelkau, J., Schaks, S., Bosse, K., Ihling, C.H., Krauth, F., Fritzsche, R., Kuhn, U., and Sinz, A. (2012). StavroX--a software for analyzing crosslinked products in protein interaction studies. Journal of the American Society for Mass Spectrometry *23*, 76-87.

Haddy, A. (2007). EPR spectroscopy of the manganese cluster of photosystem II. Photosynthesis Research *92*, 357-368.

Hall, M., Mata-Cabana, A., Akerlund, H.E., Florencio, F.J., Schroder, W.P., Lindahl, M., and Kieselbach, T. (2010). Thioredoxin targets of the plant chloroplast lumen and their implications for plastid function. Proteomics *10*, 987-1001.

Han, G., Mamedov, F., and Styring, S. (2012). Misses during water oxidation in photosystem II are S state-dependent. The Journal of Biological Chemistry *287*, 13422-13429.

Hansson, M., and Vener, A.V. (2003). Identification of three previously unknown in vivo protein phosphorylation sites in thylakoid membranes of Arabidopsis thaliana. Molecular & Cellular Proteomics : MCP *2*, 550-559.

Haumann, M., Muller, C., Liebisch, P., Iuzzolino, L., Dittmer, J., Grabolle, M., Neisius, T., Meyer-Klaucke, W., and Dau, H. (2005). Structural and oxidation state changes of the photosystem II manganese complex in four transitions of the water oxidation cycle (S0 --> S1, S1 --> S2, S2 --> S3, and S3,4 --> S0) characterized by X-ray absorption spectroscopy at 20 K and room temperature. Biochemistry *44*, 1894-1908.

Ho, F.M. (2012). Structural and mechanistic investigations of photosystem II through computational methods. Biochimica et Biophysica Acta *1817*, 106-120.

Ho, F.M., and Styring, S. (2008). Access channels and methanol binding site to the CaMn4 cluster in Photosystem II based on solvent accessibility simulations, with implications for substrate water access. Biochimica et Biophysica Acta *1777*, 140-153.

Hooft, RWW, Vriend G, Sander C, Abola EE. (1996). Errors in protein structures. *381*, 272-272.

Ido, K., Ifuku, K., Yamamoto, Y., Ishihara, S., Murakami, A., Takabe, K., Miyake, C., and Sato, F. (2009). Knockdown of the PsbP protein does not prevent assembly of the dimeric PSII core complex but impairs accumulation of photosystem II supercomplexes in tobacco. Biochimica et Biophysica Acta *1787*, 873-881.

Ido, K., Kakiuchi, S., Uno, C., Nishimura, T., Fukao, Y., Noguchi, T., Sato, F., and Ifuku, K. (2012). The conserved His-144 in the PsbP protein is important for the interaction between the PsbP N-terminus and the Cyt b559 subunit of photosystem II. The Journal of Biological Chemistry *287*, 26377-26387.

Ido, K., Nield, J., Fukao, Y., Nishimura, T., Sato, F., and Ifuku, K. (2014). Cross-linking evidence for multiple interactions of the PsbP and PsbQ proteins in a higher plant photosystem II supercomplex. The Journal of Biological Chemistry *289*, 20150-20157.

Ifuku, K. (2015). Localization and functional characterization of the extrinsic subunits of photosystem II: an update. Bioscience, Biotechnology, and Biochemistry *79*, 1223-1231.

Ifuku, K., Ishihara, S., and Sato, F. (2010). Molecular functions of oxygen-evolving complex family proteins in photosynthetic electron flow. Journal of Integrative Plant Biology *52*, 723-734.

Ifuku, K., Ishihara, S., Shimamoto, R., Ido, K., and Sato, F. (2008). Structure, function, and evolution of the PsbP protein family in higher plants. Photosynthesis Research *98*, 427-437.

Ifuku, K., Nakatsu, T., Shimamoto, R., Yamamoto, Y., Ishihara, S., Kato, H., and Sato, F. (2005a). Structure and function of the PsbP protein of photosystem II from higher plants. Photosynthesis Research *84*, 251-255.

Ifuku, K., and Sato, F. (2001). Importance of the N-terminal sequence of the extrinsic 23 kDa polypeptide in Photosystem II in ion retention in oxygen evolution. Biochimica et Biophysica Acta *1546*, 196-204.

Ifuku, K., and Sato, F. (2002). A truncated mutant of the extrinsic 23-kDa protein that absolutely requires the extrinsic 17-kDa protein for Ca2+ retention in photosystem II. Plant & Cell Physiology *43*, 1244-1249.
Ifuku, K., Yamamoto, Y., Ono, T., Ishihara, S., and Sato, F. (2005b). PsbP Protein, But Not PsbQ Protein, Is Essential for the Regulation and Stabilization of Photosystem II in Higher Plants. Plant Physiology *139*, 1175-1184.

Imre Vass, K.M.C., Zsuzsanna Deák, Steve R. mayes, James Barber (1992). Thermoluminescence and flash-oxygen characterization of the IC2 deletion mutant of Synechocystis sp. PCC 6803 lacking the Photosystem II 33 kDa protein. Biochim Biophys Acta-Bioenergetics *1102*, 195-201.

Ishihara, S., Takabayashi, A., Ido, K., Endo, T., Ifuku, K., and Sato, F. (2007). Distinct functions for the two PsbP-like proteins PPL1 and PPL2 in the chloroplast thylakoid lumen of Arabidopsis. Plant Physiol *145*, 668-679.

Ishikawa, Y., Schroder, W.P., and Funk, C. (2005). Functional analysis of the PsbP-like protein (sll1418) in Synechocystis sp. PCC 6803. Photosynthesis Research *84*, 257-262.

Jackson, S.A., and Eaton-Rye, J.J. (2015). Characterization of a Synechocystis sp. PCC 6803 double mutant lacking the CyanoP and Ycf48 proteins of Photosystem II. Photosynthesis Research *124*, 217-229.

Jackson, S.A., Hinds, M.G., and Eaton-Rye, J.J. (2012). Solution structure of CyanoP from Synechocystis sp. PCC 6803: new insights on the structural basis for functional specialization amongst PsbP family proteins. Biochimica et Biophysica Acta *1817*, 1331-1338.

Joliot, P. (1965). [Kinetics of photosynthetic fluorescence in relation to oxygen evolution]. Biochimica et Biophysica Acta *102*, 135-148.

Joliot, P. (1968). Kinetic studies of photosystem II in photosynthesis. Photochemistry and Photobiology *8*, 451-463.

Joliot, P. (2003). Period-four oscillations of the flash-induced oxygen formation in photosynthesis. Photosynthesis Research *76*, 65-72.

Joliot, P., Barbieri, G., Chabaud, R. (1969). A new model for the photochemical centers of the system II. Photosynthesis Research *10*, 309-329.

Kakiuchi, S., Uno, C., Ido, K., Nishimura, T., Noguchi, T., Ifuku, K., and Sato, F. (2012). The PsbQ protein stabilizes the functional binding of the PsbP protein to photosystem II in higher plants. Biochimica et Biophysica Acta *1817*, 1346-1351.

Kalkhof, S., and Sinz, A. (2008). Chances and pitfalls of chemical cross-linking with amine-reactive N-hydroxysuccinimide esters. Analytical and Bioanalytical Chemistry *392*, 305-312.

Kamiya, N., and Shen, J.R. (2003). Crystal structure of oxygen-evolving photosystem II from Thermosynechococcus vulcanus at 3.7-A resolution. Proceedings of the National Academy of Sciences of the United States of America *100*, 98-103.

Kashino, Y., Inoue-Kashino, N., Roose, J.L., and Pakrasi, H.B. (2006). Absence of the PsbQ protein results in destabilization of the PsbV protein and decreased oxygen evolution activity in cyanobacterial photosystem II. The Journal of Biological Chemistry *281*, 20834-20841.

Kaur, P., Kiselar, J.G., and Chance, M.R. (2009). Integrated algorithms for high-throughput examination of covalently labeled biomolecules by structural mass spectrometry. Analytical Chemistry *81*, 8141-8149.
Kern, J., Loll, B., Luneberg, C., DiFiore, D., Biesiadka, J., Irrgang, K.D., and Zouni, A. (2005).

Purification, characterisation and crystallisation of photosystem II from Thermosynechococcus elongatus cultivated in a new type of photobioreactor. Biochimica et Biophysica Acta *1706*, 147-157.

Klockenbusch, C., O'Hara, J.E., and Kast, J. (2012). Advancing formaldehyde cross-linking towards quantitative proteomic applications. Analytical and Bioanalytical Chemistry *404*, 1057-1067.

Kok, B., Forbush, B., and McGloin, M. (1970). Cooperation of charges in photosynthetic O2 evolution-I. A linear four step mechanism. Photochemistry and Photobiology *11*, 457-475.

Kopecky, V., Jr., Kohoutova, J., Lapkouski, M., Hofbauerova, K., Sovova, Z., Ettrichova, O., Gonzalez-Perez, S., Dulebo, A., Kaftan, D., Smatanova, I.K., *et al.* (2012). Raman spectroscopy adds complementary detail to the high-resolution x-ray crystal structure of photosynthetic PsbP from Spinacia oleracea. PloS One *7*, e46694.

Kosinski, J., von Appen, A., Ori, A., Karius, K., Muller, C.W., and Beck, M. (2015). Xlink Analyzer: software for analysis and visualization of cross-linking data in the context of three-dimensional structures. Journal of Structural Biology *189*, 177-183.

Krewald, V., Retegan, M., Cox, N., Messinger, J., Lubitz, W., S., D., Neesea, F., and Pantazis, D.A. (2015). Metal oxidation states in biological water splitting. Chemical Science *6*, 1676-1695.

Kuhl, H., Rogner, M., Van Breemen, J.F., and Boekema, E.J. (1999). Localization of cyanobacterial photosystem II donor-side subunits by electron microscopy and the supramolecular organization ofphotosystem II in the thylakoid membrane. European Journal of Biochemistry / FEBS *266*, 453-459.

Kulik, L.V., Epel, B., Lubitz, W., and Messinger, J. (2005). 55Mn pulse ENDOR at 34 GHz of the S0 and S2 states of the oxygen-evolving complex in photosystem II. Journal of the American Chemical Society *127*, 2392-2393.

Kulik, L.V., Epel, B., Lubitz, W., and Messinger, J. (2007). Electronic structure of the Mn4OxCa cluster in the S0 and S2 states of the oxygen-evolving complex of photosystem II based on pulse 55Mn-ENDOR and EPR spectroscopy. Journal of the American Chemical Society *129*, 13421-13435.

Kupitz, C., Basu, S., Grotjohann, I., Fromme, R., Zatsepin, N.A., Rendek, K.N., Hunter, M.S., Shoeman, R.L., White, T.A., Wang, D., *et al.* (2014). Serial time-resolved crystallography of photosystem II using a femtosecond X-ray laser. Nature *513*, 261-265.

Kuwabara, T., Murata, T., Miyao, M., and Murata, N. (1986). Partial Degradation of the 18-Kda Protein of the Photosynthetic Oxygen-Evolving Complex - a Study of a Binding-Site. Biochimica et Biophysica Acta *850*, 146-155.

Laskoswki RA, MacArthur MW, Moss DS & Thorton JM. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Cryst. 26, 283-291.

Leuschner, C., and Bricker, T.M. (1996). Interaction of the 33 kDa extrinsic protein with photosystem II: rebinding of the 33 kDa extrinsic protein to photosystem II membranes which contain four, two, or zero manganese per photosystem II reaction center. Biochemistry *35*, 4551-4557.

Liu, H., Weisz, D.A., and Pakrasi, H.B. (2015). Multiple copies of the PsbQ protein in a cyanobacterial photosystem II assembly intermediate complex. Photosynthesis Research *126*, 375-383.

Liu, H., Zhang, H., Weisza, D.A., Vidavsky, I., Gross, M.L., and Pakrasi, H.B. (2014). MS-based cross-linking analysis reveals the location of the PsbQ protein in cyanobacterial photosystem II. Proceedings of the National Academy of Sciences of the United States of America *111*, 4638-4643.

Liu, J., Yang, H., Lu, Q., Wen, X., Chen, F., Peng, L., Zhang, L., and Lu, C. (2012). PsbP-domain protein1, a nuclear-encoded thylakoid lumenal protein, is essential for photosystem I assembly in Arabidopsis. The Plant Cell *24*, 4992-5006.

Liu., H., Frankel, L.K., and Bricker, T.M. (2007). Functional analysis of photosystem II in a PsbO-1-deficient mutant in Arabidopsis thaliana. Biochemistry *46*, 7607-7613.

Ljungberg, U., Jansson, C., Andersson, B., and Akerlund, H.E. (1983). Reconstitution of oxygen evolution in high salt washed photosystem II particles. Biochemical and Biophysical Research Communications *113*, 738-744.

Loll, B., Kern, J., Saenger, W., Zouni, A., and Biesiadka, J. (2005). Towards complete cofactor arrangement in the 3.0 A resolution structure of photosystem II. Nature *438*, 1040-1044.

Lorch, S., Capponi, S., Pieront, F., and Bondar, A.N. (2015). Dynamic Carboxylate/Water Networks on the Surface of the PsbO Subunit of Photosystem II. The Journal of Physical Chemistry. B *119*, 12172-12181.

Luo, H., Jackson, S.A., Fagerlund, R.D., Summerfield, T.C., and Eaton-Rye, J.J. (2014). The importance of the hydrophilic region of PsbL for the plastoquinone electron acceptor complex of Photosystem II. Biochimica et Biophysica Acta *1837*, 1435-1446.

Madler, S., Bich, C., Touboul, D., and Zenobi, R. (2009). Chemical cross-linking with NHS esters: a systematic study on amino acid reactivities. Journal of Mass Spectrometry *44*, 694-706.

Mavankal, G., Mccain, D.C., and Bricker, T.M. (1986). Effects of Chloride on Paramagnetic Coupling of Manganese in Calcium Chloride-Washed Photosystem-Ii Preparations. FEBS Letters *202*, 235-239.

Mayfield, S.P., Bennoun, P., and Rochaix, J.D. (1987a). Expression of the nuclear encoded OEE1 protein is required for oxygen evolution and stability of photosystem II particles in Chlamydomonas reinhardtii. The EMBO Journal *6*, 313-318.

Mayfield, S.P., Rahire, M., Frank, G., Zuber, H., and Rochaix, J.D. (1987b). Expression of the nuclear gene encoding oxygen-evolving enhancer protein 2 is required for high levels of photosynthetic oxygen evolution in Chlamydomonas reinhardtii. Proceedings of the National Academy of Sciences of the United States of America *84*, 749-753.

McEvoy, J.P., and Brudvig, G.W. (2006). Water-splitting chemistry of photosystem II. Chem Rev *106*, 4455-4483.

Meades, G.D., Jr., McLachlan, A., Sallans, L., Limbach, P.A., Frankel, L.K., and Bricker, T.M. (2005). Association of the 17-kDa extrinsic protein with photosystem II in higher plants. Biochemistry *44*, 15216-15221.

Michoux, F., Boehm, M., Bialek, W., Takasaka, K., Maghlaoui, K., Barber, J., Murray, J.W., and Nixon, P.J. (2014). Crystal structure of CyanoQ from the thermophilic cyanobacterium Thermosynechococcus elongatus and detection in isolated photosystem II complexes. Photosynthesis Research *122*, 57-67.

Migneault, I., Dartiguenave, C., Bertrand, M.J., and Waldron, K.C. (2004). Glutaraldehyde: behavior in aqueous solution, reaction with proteins, and application to enzyme crosslinking. BioTechniques *37*, 790-796, 798-802.

Miyao, M., and Murata, N. (1983). Partial disintegration and reconstitution of the photosynthetic oxygen evolution system. Binding of 24 kilodalton and 18 kilodalton polypeptides. Biochimica et Biophysica Acta (BBA) - Bioenergetics *725*, 87-93.

Miyao, M., and Murata, N. (1984). Calcium ions can be substituted for the 24-kDa polypeptide in photosynthetic oxygen evolution. FEBS Letters *168*, 118-120.

Miyao, M., and Murata, N. (1985). The Cl− effect on photosynthetic oxygen evolution: interaction of Cl− with 18-kDa, 24-kDa and 33-kDa proteins. FEBS Letters *180*, 303-308.

Miyao, M., Murata, N. (1984). Role of the 33- kDa polypeptide in preserving Mn in the photosynthetic oxygen- evolution system and its replacement by chloride ions. FEBS Letters *170*, 350-354.

Miyao, M., Murata, N., Lavorel, J., Maison-Peteri, B., Boussac, A., Etienne, A. (1987). Effect of the 33-kDa protein on the S-state transitions in photosynthetic oxygen evolution. Biochimica et Biophysica Acta (BBA) - Bioenergetics *890*, 151-159.

Mullineaux, C.W. (1999). The thylakoid membranes of cyanobacteria: structure, dynamics and function. Aust J Plant Physiol *26*, 671-677.

Mummadisetti, M.P., Frankel, L.K., Bellamy, H.D., Sallans, L., Goettert, J.S., Brylinski, M., Limbach, P.A., and Bricker, T.M. (2014). Use of protein cross-linking and radiolytic footprinting to elucidate PsbP and PsbQ interactions within higher plant Photosystem II. Proceedings of the National Academy of Sciences of the United States of America *111*, 16178-16183.

Murakami, R., Ifuku, K., Takabayashi, A., Shikanai, T., Endo, T., and Sato, F. (2002). Characterization of an Arabidopsis thaliana mutant with impaired psbO, one of two genes encoding extrinsic 33-kDa proteins in photosystem II. FEBS Letters *523*, 138-142.

Murakami, R., Ifuku, K., Takabayashi, A., Shikanai, T., Endo, T., and Sato, F. (2005). Functional dissection of two Arabidopsis PsbO proteins: PsbO1 and PsbO2. The FEBS Journal *272*, 2165-2175.

Murray, J.W., and Barber, J. (2007). Structural characteristics of channels and pathways in photosystem II including the identification of an oxygen channel. Journal of Structural Biology *159*, 228-237.

Munekage, Y., Hashimoto, M., Miyake, C., Tomizawa, K., Endo, T., Tasaka, M., and Shikanai, T. (2004). Cyclic electron flow around photosystem I is essential for photosynthesis. Nature *429*, 579-582.

Nagao, R., Suzuki, T., Okumura, A., Niikura, A., Iwai, M., Dohmae, N., Tomo, T., Shen, J.R., Ikeuchi, M., and Enami, I. (2010). Topological analysis of the extrinsic PsbO, PsbP and PsbQ proteins in a green algal PSII complex by cross-linking with a water-soluble carbodiimide. Plant & Cell Physiology *51*, 718-727.

Nakatani, H.Y. (1984). Photosynthetic oxygen evolution does not require the participation of polypeptides of 16 and 24 kilodaltons. Biochemical and Biophysical Research Communications *120*, 299-304.

Nishimura, T., Nagao, R., Noguchi, T., Nield, J., Sato, F., and Ifuku, K. (2016). The N-terminal sequence of the extrinsic PsbP protein modulates the redox potential of Cyt b559 in photosystem II. Scientific Reports *6*, 21490.

Noguchi, T. (2008a). Fourier transform infrared analysis of the photosynthetic oxygen-evolving center. Coordination Chemistry Reviews *252*, 336-346.

Noguchi, T. (2008b). FTIR detection of water reactions in the oxygen-evolving centre of photosystem II. Philosophical transactions of the Royal Society of London. Series B, Biological Sciences *363*, 1189-1194; discussion 1194-1185.

Noguchi, T. (2015). Fourier transform infrared difference and time-resolved infrared detection of the electron and proton transfer dynamics in photosynthetic water oxidation. BBA Bioenergetics *1847*, 35-45.

Offenbacher, A.R., Polander, B.C., and Barry, B.A. (2013). An intrinsically disordered photosystem II subunit, PsbO, provides a structural template and a sensor of the hydrogen-bonding network in photosynthetic water oxidation. The Journal of Biological Chemistry *288*, 29056-29068.

Okumura, A., Nagao, R., Suzuki, T., Yamagoe, S., Iwai, M., Nakazato, K., and Enami, I. (2008). A novel protein in Photosystem II of a diatom Chaetoceros gracilis is one of the extrinsic proteins located on lumenal side and directly associates with PSII core components. Biochimica et Biophysica Acta *1777*, 1545-1551.

Ono, T., Izawa, S., and Inoue, Y. (1992). Structural and functional modulation of the manganese cluster in Ca(2+)-depleted photosystem II induced by binding of the 24-kilodalton extrinsic protein. Biochemistry *31*, 7648-7655.

Pearson, W.R. (2013). An introduction to sequence similarity ("homology") searching. Current protocols in bioinformatics / editorial board, Andreas D. Baxevanis ... [et al.] *Chapter 3*, Unit3 1.

Peltier, J.B., Emanuelsson, O., Kalume, D.E., Ytterberg, J., Friso, G., Rudella, A., Liberles, D.A., Soderberg, L., Roepstorff, P., von Heijne, G., *et al.* (2002). Central functions of the lumenal and peripheral thylakoid proteome of Arabidopsis determined by experimentation and genome-wide prediction. The Plant Cell *14*, 211-236.

Peri, S., Steen, H., and Pandey, A. (2001). GPMAW--a software tool for analyzing proteins and peptides. Trends in Biochemical Sciences *26*, 687-689.

Perkins, D.N., Pappin, D.J., Creasy, D.M., and Cottrell, J.S. (1999). Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis *20*, 3551-3567.

Petrek, M., Otyepka, M., Banas, P., Kosinova, P., Koca, J., and Damborsky, J. (2006). CAVER: a new tool to explore routes from protein clefts, pockets and cavities. BMC Bioinformatics *7*, 316.

Phuthong, W., Huang, Z., Wittkopp, T.M., Sznee, K., Heinnickel, M.L., Dekker, J.P., Frese, R.N., Prinz, F.B., and Grossman, A.R. (2015). The Use of Contact Mode Atomic Force Microscopy in Aqueous Medium for Structural Analysis of Spinach Photosynthetic Complexes. Plant Physiology *169*, 1318-1332.

Popelka, H., and Yocum, C. (2012). Probing the N-terminal sequence of spinach PsbO: evidence that essential threonine residues bind to different functional sites in eukaryotic photosystem II. Photosynthesis Research *112*, 117-128.

Popelkova, H., Commet, A., Kuntzleman, T., and Yocum, C.F. (2008). Inorganic cofactor stabilization and retention: the unique functions of the two PsbO subunits of eukaryotic photosystem II. Biochemistry *47*, 12593-12600.

Popelkova, H., Im, M.M., D'Auria, J., Betts, S.D., Lydakis-Simantiris, N., and Yocum, C.F. (2002a). N-terminus of the photosystem II manganese stabilizing protein: effects of sequence elongation and truncation. Biochemistry *41*, 2702-2711.

Popelkova, H., Im, M.M., and Yocum, C.F. (2002b). N-terminal truncations of manganese stabilizing protein identify two amino acid sequences required for binding of the eukaryotic protein to photosystem II and reveal the absence of one binding-related sequence in cyanobacteria. Biochemistry *41*, 10038-10045.

Popelkova, H., Im, M.M., and Yocum, C.F. (2003). Binding of manganese stabilizing protein to photosystem II: identification of essential N-terminal threonine residues and domains that prevent nonspecific binding. Biochemistry *42*, 6193-6200.

Pospisil, P. (2009). Production of reactive oxygen species by photosystem II. Biochimica et Biophysica Acta *1787*, 1151-1160.

Pospisil, P. (2014). The role of metals in production and scavenging of reactive oxygen species in photosystem II. Plant & Cell Physiology *55*, 1224-1232.

Pospísil, P., Arató, A., Krieger-Liszkay, A., and Rutherford, A.W. (2004). Hydroxyl radical generation by photosystem II. Biochemistry *43*, 6783-6792.

Pushkar, Y., Yano, J., Sauer, K., Boussac, A., and Yachandra, V.K. (2008). Structural changes in the Mn4Ca cluster and the mechanism of photosynthetic water splitting. Proceedings of the National Academy of Sciences of the United States of America *105*, 1879-1884.

Puthiyaveetil, S., and Kirchhoff, H. (2013). A phosphorylation map of the photosystem II supercomplex C2S2M2. Frontiers in Plant Science *4*, 459.

Rathner, P., Rathner, A., Hornicakova, M., Wohlschlager, C., Chandra, K., Kohoutova, J., Ettrich, R., Wimmer, R., and Muller, N. (2015). Solution NMR and molecular dynamics reveal a persistent alpha helix within the dynamic region of PsbQ from photosystem II of higher plants. Proteins *83*, 1677-1686.

Reiland, S., Messerli, G., Baerenfaller, K., Gerrits, B., Endler, A., Grossmann, J., Gruissem, W., and Baginsky, S. (2009). Large-scale Arabidopsis phosphoproteome profiling reveals novel chloroplast kinase substrates and phosphorylation networks. Plant Physiol *150*, 889-903.

Roose, J.L., Frankel, L.K., and Bricker, T.M. (2010). Documentation of significant electron transport defects on the reducing side of photosystem II upon removal of the PsbP and PsbQ extrinsic proteins. Biochemistry *49*, 36-41.

Roose, J.L., Frankel, L.K., and Bricker, T.M. (2011a). Developmental defects in mutants of the PsbP domain protein 5 in Arabidopsis thaliana. PloS One *6*, e28624.

Roose, J.L., Frankel, L.K., and Bricker, T.M. (2014). The PsbP domain protein 1 functions in the assembly of lumenal domains in photosystem I. The Journal of Biological Chemistry *289*, 23776-23785.

Roose, J.L., Frankel, L.K., Mummadisetti, M.P., and Bricker, T.M. (2016). The extrinsic proteins of photosystem II: update. Planta.

Roose, J.L., Kashino, Y., and Pakrasi, H.B. (2007a). The PsbQ protein defines cyanobacterial Photosystem II complexes with highest activity and stability. Proceedings of the National Academy of Sciences of the United States of America *104*, 2548-2553.

Roose, J.L., Wegener, K.M., and Pakrasi, H.B. (2007b). The extrinsic proteins of Photosystem II. Photosynthesis Research *92*, 369-387.

Roose, J.L., Yocum, C.F., and Popelkova, H. (2011b). Binding stoichiometry and affinity of the manganese-stabilizing protein affects redox reactions on the oxidizing side of photosystem II. Biochemistry *50*, 5988-5998.

Saito, K., Rutherford, A.W., and Ishikita, H. (2013a). Mechanism of proton-coupled quinone reduction in Photosystem II. Proceedings of the National Academy of Sciences of the United States of America *110*, 954-959.

Saito, K., Rutherford, A.W., and Ishikita, H. (2013b). Mechanism of tyrosine D oxidation in Photosystem II. Proceedings of the National Academy of Sciences of the United States of America *110*, 7690-7695.
Sato, N. (2010). Phylogenomic and structural modeling analyses of the PsbP superfamily reveal multiple small segment additions in the evolution of photosystem II-associated PsbP protein in green plants. Molecular Phylogenetics and Evolution *56*, 176-186.

Sauer, K., Yano, J., and Yachandra, V.K. (2008). X-Ray spectroscopy of the photosynthetic oxygen-evolving complex. Coordination Chemistry Reviews *252*, 318-335.

Schubert, M., Petersson, U.A., Haas, B.J., Funk, C., Schroder, W.P., and Kieselbach, T. (2002). Proteome map of the chloroplast lumen of Arabidopsis thaliana. The Journal of Biological Chemistry *277*, 8354-8365.
Scigelova, M., Hornshaw, M., Giannakopulos, A., and Makarov, A. (2011). Fourier transform mass spectrometry. Molecular & Cellular Proteomics : MCP *10*, M111 009431.

Sharma, J., Panico, M., Barber, J., and Morris, H.R. (1997). Purification and determination of intact molecular mass by electrospray ionization mass spectrometry of the photosystem II reaction center subunits. The Journal of Biological Chemistry *272*, 33153-33157.

Shen, J.R. (1997). Possible functional differences between dimer and monomer of Photosystem II complex (Kluwer Academic Publishers).

Shen, J.R. (2015). The Structure of Photosystem II and the Mechanism of Water Oxidation in Photosynthesis. Annual Review of Plant Biology *66*, 23-48.

Shukla, A.K., and Futrell, J.H. (2000). Tandem mass spectrometry: dissociation of ions by collisional activation. Journal of Mass Spectrometry : JMS *35*, 1069-1090.

Shutova, T., Klimov, V.V., Andersson, B., and Samuelsson, G. (2007). A cluster of carboxylic groups in PsbO protein is involved in proton transfer from the water oxidizing complex of Photosystem II. Biochimica et Biophysica Acta *1767*, 434-440.

Sleno, L., and Volmer, D.A. (2004). Ion activation methods for tandem mass spectrometry. Journal of Mass Spectrometry : JMS *39*, 1091-1112.

Sriswasdi, S., Harper, S.L., Tang, H.Y., and Speicher, D.W. (2014). Enhanced identification of zero-length chemical cross-links using label-free quantitation and high-resolution fragment ion spectra. Journal of Proteome Research *13*, 898-914.

Suorsa, M., Regel, R.E., Paakkarinen, V., Battchikova, N., Herrmann, R.G., and Aro, E.M. (2004). Protein assembly of photosystem II and accumulation of subcomplexes in the absence of low molecular mass subunits PsbL and PsbJ. European Journal of Biochemistry / FEBS *271*, 96-107.

Sutherland, B.W., Toews, J., and Kast, J. (2008). Utility of formaldehyde cross-linking and mass spectrometry in the study of protein-protein interactions. Journal of Mass Spectrometry : JMS *43*, 699-715.
Taka-aki Ono, Y.I. (1985). S-state turnover in the O2-evolving system of CaCl2-washed Photosystem II particles depleted of three peripheral proteins as measured by thermoluminescence. Removal of 33 kDa protein inhibits S3 to S4 transition. BBA Bioenergetics *806*, 331-340.

Takamoto, K., and Chance, M.R. (2006). Radiolytic protein footprinting with mass spectrometry to probe the structure of macromolecular complexes. Annual Review of Biophysics and Biomolecular Structure *35*, 251-276.

Thornton, L.E., Ohkawa, H., Roose, J.L., Kashino, Y., Keren, N., and Pakrasi, H.B. (2004). Homologs of plant PsbP and PsbQ proteins are necessary for regulation of photosystem ii activity in the cyanobacterium Synechocystis 6803. The Plant Cell *16*, 2164-2175.

Tikkanen, M., and Aro, E.M. (2012). Thylakoid protein phosphorylation in dynamic regulation of photosystem II in higher plants. Biochimica et Biophysica Acta *1817*, 232-238.

Toews, J., Rogalski, J.C., Clark, T.J., and Kast, J. (2008). Mass spectrometric identification of formaldehyde-induced peptide modifications under in vivo protein cross-linking conditions. Analytica Chimica Acta *618*, 168-183.

Tomita, M., Ifuku, K., Sato, F., and Noguchi, T. (2009). FTIR evidence that the PsbP extrinsic protein induces protein conformational changes around the oxygen-evolving Mn cluster in photosystem II. Biochemistry *48*, 6318-6325.

Turkina, M.V., Kargul, J., Blanco-Rivero, A., Villarejo, A., Barber, J., and Vener, A.V. (2006). Environmentally modulated phosphoproteome of photosynthetic membranes in the green alga Chlamydomonas reinhardtii. Molecular & Cellular Proteomics : MCP *5*, 1412-1425.

Umena, Y., Kawakami, K., Shen, J.R., and Kamiya, N. (2011). Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 A. Nature *473*, 55-60.

Vassiliev, S., Comte, P., Mahboob, A., and Bruce, D. (2010). Tracking the flow of water through photosystem II using molecular dynamics and streamline tracing. Biochemistry *49*, 1873-1881.

Veerman, J., Bentley, F.K., Eaton-Rye, J.J., Mullineaux, C.W., Vasil'ev, S., and Bruce, D. (2005). The PsbU subunit of photosystem II stabilizes energy transfer and primary photochemistry in the phycobilisome-photosystem II assembly of Synechocystis sp. PCC 6803. Biochemistry *44*, 16939-16948.

Vener, A.V. (2007). Environmentally modulated phosphorylation and dynamics of proteins in photosynthetic membranes. Biochimica et Biophysica Acta *1767*, 449–457.

Vener, A.V., Harms, A., Sussman, M.R., and Vierstra, R.D. (2001). Mass spectrometric resolution of reversible protein phosphorylation in photosynthetic membranes of Arabidopsis thaliana. The Journal of Biological Chemistry *276*, 6959-6966.

Vyas, V.K., Ukawala, R.D., Ghate, M., and Chintha, C. (2012). Homology modeling a fast tool for drug discovery: current perspectives. Indian Journal of Pharmaceutical Sciences *74*, 1-17.

Watanabe, M., Iwa, M., Narikawa, R., and Ikeuchi, M. (2009). Is the Photosystem II Complex a Monomer or a Dimer? . Plant Cell Physiology *50*, 1674-1680.

Xu, H., and Freitas, M.A. (2007). A mass accuracy sensitive probability based scoring algorithm for database searching of tandem mass spectrometry data. BMC Bioinformatics *8*, 133.Xu, H., and Freitas, M.A. (2008). Monte carlo simulation-based algorithms for analysis of shotgun proteomic data. Journal of Proteome Research *7*, 2605-2615.

Xu, H., and Freitas, M.A. (2009). MassMatrix: a database search program for rapid characterization of proteins and peptides from tandem mass spectrometry data. Proteomics *9*, 1548-1555.

Yamamoto, Y., Nakayama, S., Cohn, C.L., and Krogmann, D.W. (1987). Highly efficient purification of the 33-, 24-, and 18-kDa proteins in spinach photosystem II by butanol/water phase partitioning and high-performance liquid chromatography. Archives of Biochemistry and Biophysics *255*, 156-161.

Yang, B., Wu, Y.J., Zhu, M., Fan, S.B., Lin, J., Zhang, K., Li, S., Chi, H., Li, Y.X., Chen, H.F., *et al.* (2012). Identification of cross-linked peptides from complex samples. Nature Methods *9*, 904-906.

Yano, J., Kern, J., Sauer, K., Latimer, M.J., Pushkar, Y., Biesiadka, J., Loll, B., Saenger, W., Messinger, J., Zouni, A., *et al.* (2006). Where water is oxidized to dioxygen: structure of the photosynthetic Mn4Ca cluster. Science *314*, 821-825.

Yano, J., Kern, J., Sauer, K., Latimer, M. J., Pushkar, Y., et al. (2006). Where water is oxidized to dioxygen: structure of the photosynthetic Mn4Ca cluster. Science *314*, 821-825.

Yano, J., and Yachandra, V.K. (2008). Where water is oxidized to dioxygen: structure of the photosynthetic Mn4Ca cluster from X-ray spectroscopy. Inorganic Chemistry 47, 1711-1726.

Yee, A.A., Savchenko, A., Ignachenko, A., Lukin, J., Xu, X.H., Skarina, T., Evdokimova, E., Liu, C.S., Semesi, A., Guido, V., *et al.* (2005). NMR and x-ray crystallography, complementary tools in structural proteomics of small proteins. Journal of the American Chemical Society *127*, 16512-16517.

Yi, X., Hargett, S.R., Liu, H., Frankel, L.K., and Bricker, T.M. (2007). The PsbP protein is required for photosystem II complex assembly/stability and photoautotrophy in Arabidopsis thaliana. The Journal of Biological Chemistry *282*, 24833-24841.

Zubrzycki, I.Z., Frankel, L.K., Russo, P.S., and Bricker, T.M. (1998). Hydrodynamic studies on the manganese-stabilizing protein of photosystem II. Biochemistry *37*, 13553-13558.

CHAPTER 2

MATERIALS AND METHODS

2.1 PSII Preparations

The chloroplasts were isolated from market spinach (*Spinacea oleracea*) leaves. Spinach is the best source to obtain photosynthetically active chloroplasts at high yield. The de-veined spinach were ground in a blender using chloroplast isolation buffer (100 mM sucrose, 200 mM NaCl, 5 mM $MgCl_2$ and 50mM Na-K phosphate, pH 7.4). After grinding, the solution was passed through several layers of cheesecloth to remove big chunks of spinach. The filtered solution containing broken cells was centrifuged at 1400 x g for 5 min, to separate the debris. The pellets thus obtained were suspended in grinding buffer or first-resuspension buffer (50 mM Mes-NaOH, pH 6.0, 300 mM sucrose). This chloroplast suspension was used for chlorophyll measurements by the method of Arnon (Arnon, 1949).

The chlorophyll assays were performed using 5-20 uL chloroplasts completely resuspended in 1 mL of 80% methanol, and spun down in a micro-centrifuge at 14,000 x g for 5 minutes. The supernatants were collected in a fresh vial and read at 645 nm, 663 nm and 710 nm on a Shimadzu UV-spectrophotometer. The following equation was used to calculate the chlorophyll concentration (Lichtenthaler, 1987).

$$[\text{Chl}]\ \text{mg/mL} = \frac{((20.2*(A645-A710) + (8.02*(A663-A710))}{10} \text{ X } \frac{\text{mL of 80\% acetone}}{\mu\text{L of sample}}$$

The chl a/b ratio of the chloroplasts is generally in a range of 2.5 - 3.2 at this stage. The extracted chloroplasts were brought to 2-3 mg/mL chlorophyll in the resuspension buffer. Sub-chloroplast membranes containing mostly PSII, and capable of higher oxygen evolution rates, were isolated by treating the chloroplasts with a non-ionic detergent, Triton-X-100 and incubated at 4 ˚C for 30 min

in the dark. These samples were spun down at 5,000 g to separate the heavy and light material (the sedimented soft pellets contain any large chunks from chloroplasts extraction and detergent). The supernatants were spun down further at 38,720 x g (or 18,000 rpm in SS-34 Sorvall rotor) for 25 minutes. The pellets were washed and resuspended again in the resuspension buffer and the previous slow (5,000 g) and fast spin (18,000 g) process was repeated to obtain high quality of oxygen evolving membranes, mostly containing PSII. The final pellets were suspended in 50 mM Mes-NaOH, pH 6.0, 300 mM sucrose and 15 mM NaCl (SMN) buffer and their chlorophyll measurements were performed. A good PSII membrane preparation has a/b chlorophyll ratio in a range of 1.8-2.0. These membranes are suspended in SMN buffer at 2 mg/mL chl and frozen as aliquots until use.

Oxygen evolution rates were measured from freshly extracted PSII membranes on a Hansatech-oxygen electrode. In every PSII preparation, the oxygen evolution rates were >400 m moles $O_2 \cdot mg\ chl^{-1} \cdot h^{-1}$. Protein crosslinking experiments were performed using BS3 (ProteoChem, Inc.) and EDC (ThermoFisher Scientific) crosslinkers.

2.2 Protein Crosslinking Procedures

In this dissertation two crosslinkers, BS3 and EDC are used. Protein crosslinkers form an amide bond between two residues, separated by a distance consistent with the arm-length of a protein crosslinker. BS3 is a homobifunctional, water-soluble, membrane impermeable and a non-cleavable crosslinker that crosslinks different residues based on the pH of the medium. It has N-hydroxysulfosuccinimide (NHS) ester as its head groups on either ends of an 8-carbon spacer arm, thus it can crosslink residues that are upto 11.4 Å apart (Staros, 1982). BS3 crosslinker can crosslink primary amino groups (seen at the N-terminus of a protein and the ε-amino group of lysines) at pH 7.0 – 9.0. However, at low pH, it can also crosslink threonyl-, tyrosyl-, and seryl- residues, (Kalkhof and Sinz, 2008b; Madler et al., 2009; Sinz, 2006; Swaim et al., 2004). The crosslinking reaction is

shown in the FIGURE 2.1. For BS3 crosslinking reaction, PS II membranes were suspended at a chl concentration of 100 μg/mL in SMN buffer and treated with various concentrations of BS3 (0–10 mM) for 1 h at room temperature and in the dark for BS3 crosslinking experiments. Bringing the reaction mixture to 30 mM ammonium bicarbonate and incubating it for 20 min at room temperature quenched the reaction. The reaction contents were spun down at 38,000 x g for 25 minutes and the pellets were uniformly suspended using a homogenizer in the SMN buffer containing 1.0 M NaCl, and this solution was left on ice, in the dark, for one hour. The high salt concentration can break the interactions of the proteins PsbP and PsbQ with PSII, and release these proteins in the buffer medium. This solution was again spun down at 38,000 g for 25 minutes, in which the supernatants contained PsbP, PsbQ, and any of their crosslinked products (not crosslinked to the membrane proteins or PsbO. The pellets were washed to remove excess salt, and frozen at -80 ˚C. The supernatants were dialyzed overnight against 10 mM Mes, pH 6.0 using, a 6–8-kDa membrane (Spectrum Laboratories, Inc.), centrifuged for 25 min at 39,000 × g, and were concentrated by ultrafiltration using a 10-kDa cutoff membrane (Millipore Co.). Protein concentrations were determined using the BCA protein assay (Smith et al., 1985).

EDC is a water-soluble carbodiimide that activates carboxyl residues to form a stable amide bond with primary amines. Hence, it can crosslink the N- to C-termini of a protein, C-terminal carboxyl to Lys, Asp to Lys and Glu to Lys. It is a zero-length crosslinker that crosslinks the residues within a van der Waals distance. Sulfo- N-hydroxysulfosuccinimide (NHS) increases the reaction efficiency of EDC crosslinking reaction by forming a stable intermediate with o-acylisourea (formed on reaction of EDC with reactive carboxyl-group) for reaction with primary amines later.

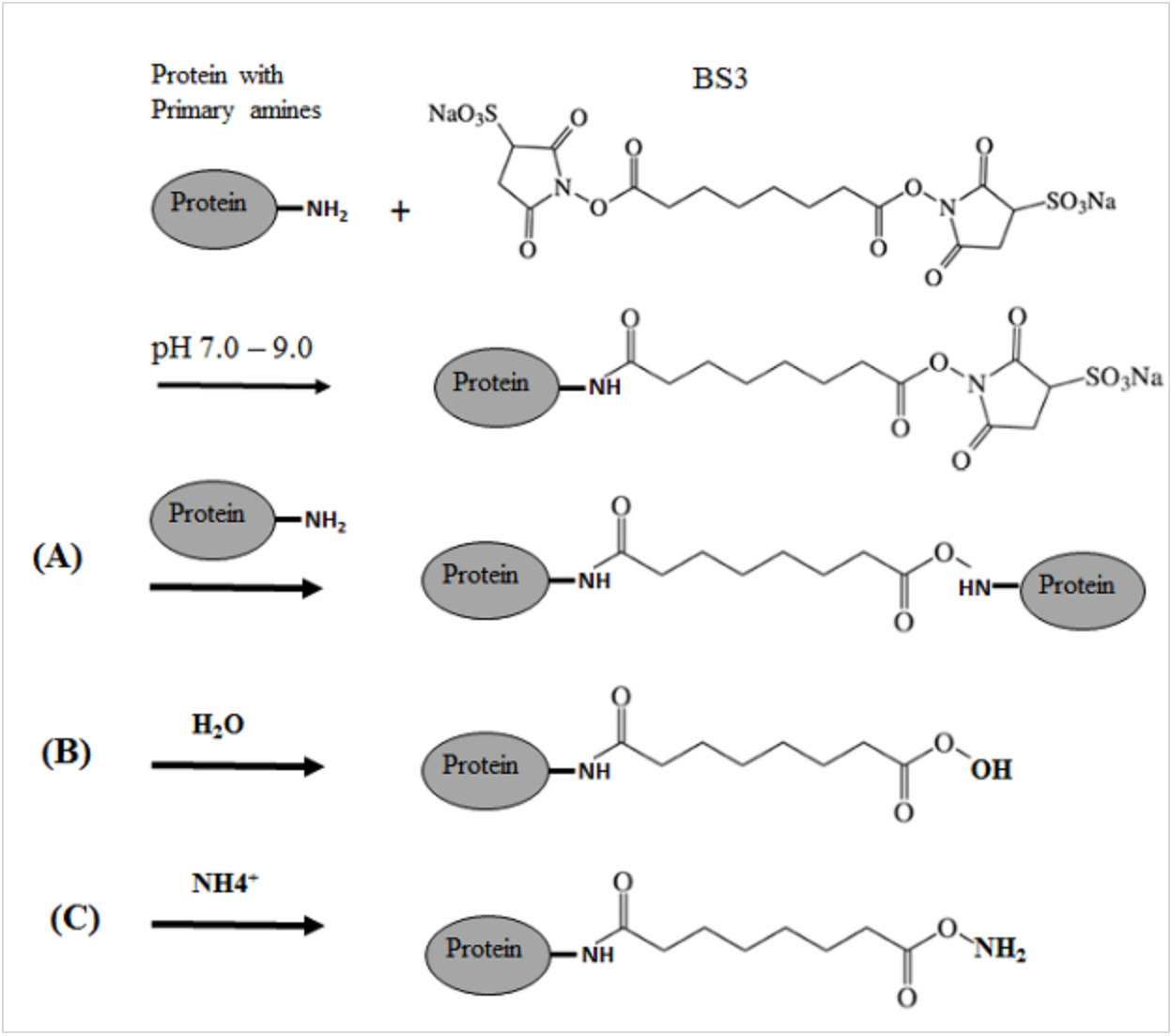
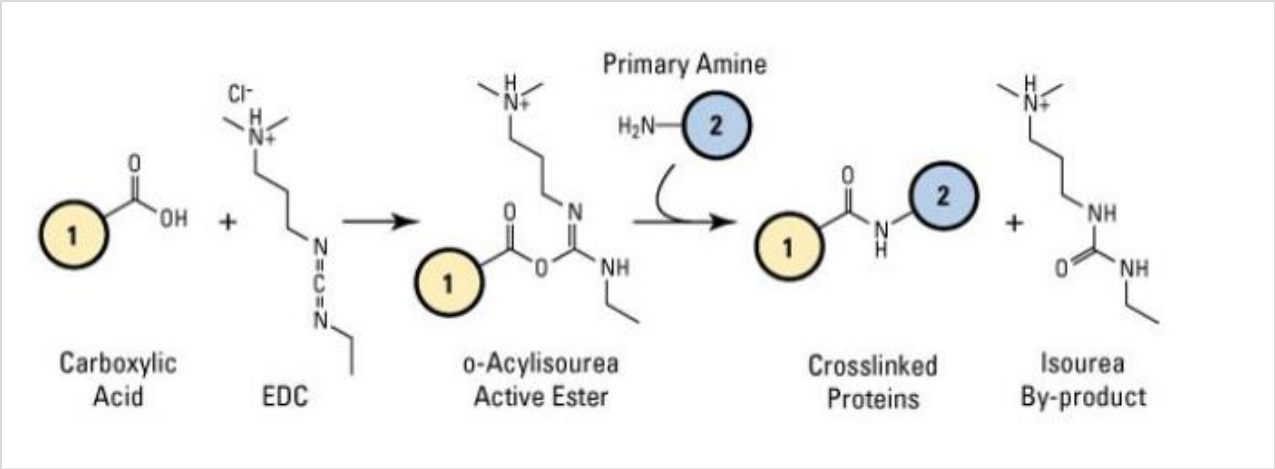


FIGURE 2.1. The Crosslinking of Proteins with BS3.

Primary amines e.g. at the N-terminus of a protein, or a ε-amino group of lysine react at basic pH (7.0 – 9.0). Tyrosine, Serine and Threonine (hydroxy-group containing residues) also react with the crosslinker, however, at acidic pH (6.0 ≤ 7.0).
A. Primary amine of a protein reacts with BS3 to release N-hydroxysulfosuccinamide group. **B.** Reaction with water. **C.** Reaction with ammonium ion or ammonium bicarbonate (usually used as a reaction quencher).

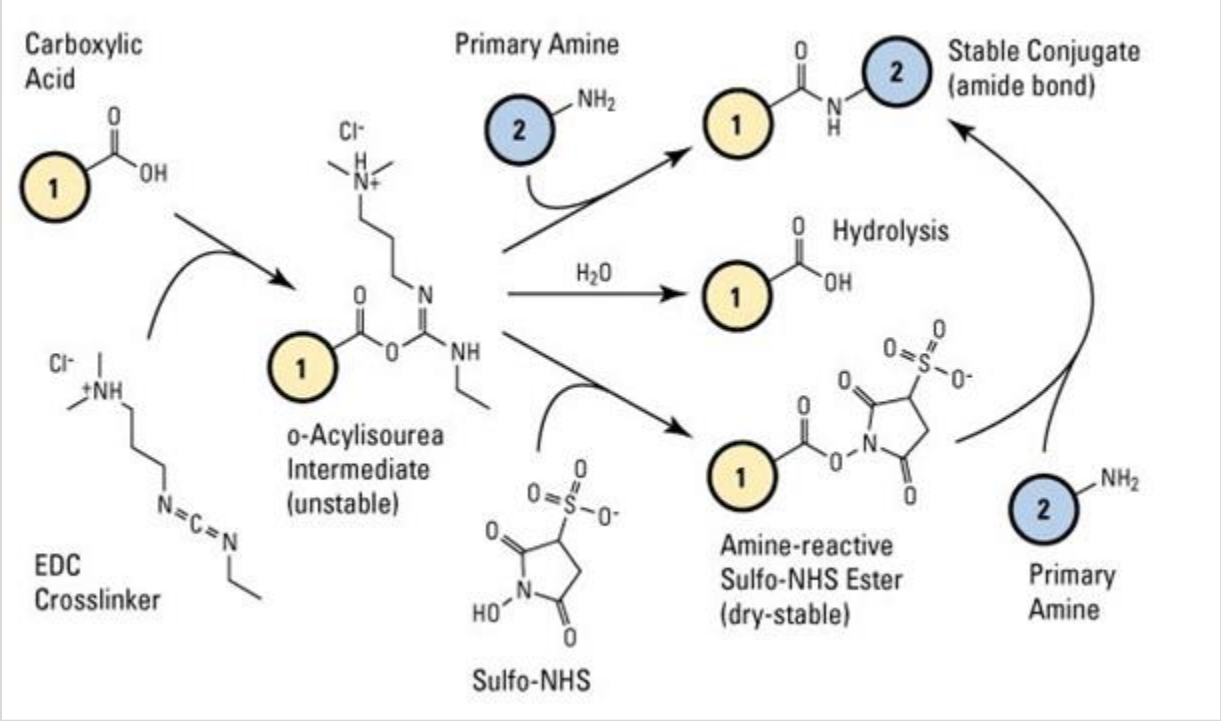
For protein-crosslinking with EDC, the oxygen-evolving PSII membranes were suspended at 200 ug/mL in the crosslinking buffer containing 25% glycerol, 10mM $MgCl_2$, 5mM $CaCl_2$ and 50mM Mes, pH 6.0. These membranes were treated with 6.25 mM EDC and 5mM sulfo-NHS. The crosslinking chemistry is shown in FIGURE 2.2 and 2.3. The reaction was quenched with 100mM



Thermofisher.com

FIGURE 2.2. The EDC Crosslinking Reaction Scheme.

Molecules 1 and 2 are different proteins in our experiments (although these can be peptides or any chemicals with carboxylate and primary amine groups). EDC reaction with carboxylic acids give o-acylisourea intermediates, which readily reacts with primary amines to form isourea, and thus forming an amide bond between the carboxylates and amine groups.



Thermofisher.com

FIGURE 2.3. The Sulfo-NHS Reaction Scheme.

Sulfo-NHS increases the half-life of reaction intermediate o-acylisourea, which is easily hydrolysable. It forms another intermediate which is a Sulfo-NHS Ester, which can react with primary amines, to release Sulfo-NHS again.

ammonium bicarbonate. The reaction contents from EDC crosslinking reaction were harvested by centrifugation for 25 min at 39,000 × g, and the final pellets were resuspended in 1.0 M NaCl (similar to BS3 procedure) for 1 h at 4 °C to release any PsbP and PsbQ proteins not crosslinked to intrinsic membrane protein or PsbO. Supernatants were collected and the PS II membranes were pelleted by centrifugation for 25 min at 39,000 × g, and washed twice in 10mM Mes, pH 6.0, to remove excess NaCl used for washing the membranes. The membranes free of NaCl were then washed with $CaCl_2$ to release free PsbO, and any crosslinked products with PsbP and PsbQ. The membranes were again pelleted and washed as described above, resuspended at 2mg per mL Chl concentration in SMN buffer and stored in -80º freezer. Supernatants, from 1M NaCl wash contained free PsbP, free PsbQ, and PsbP–PsbQ crosslinked products. And the supernatants from 1M $CaCl_2$ treatments contained free PsbO and its crosslinked products with PsbP and PsbQ. These salt-extracted samples were dialyzed overnight against 10 mM Mes-NaOH, pH 6.0.

2.3 Electrophoresis and Protein Digestion

For the crosslinking experiments, the protein samples were resolved on a standard acrylamide 12.5–20% LiDS– PAGE gradient gel (Delepelaire and Chua, 1979). After electrophoresis, the gels were stained with Coomassie Blue, destained, and the protein bands of interest were excised out of the gel. The extracted gel-fragments were then digested with different proteases (trypsin or trypsin + Lys-C). In some cases, the mass spectrometry-compatible ProteaseMax (Promega) was included during digestion.

2.4 In-gel Protease Digestion with Trypsin and Extraction of Tryptic Fragments

The protein bands of interest were excised from the Coomassie blue stained-, followed by destained- gel (methanol + acetic acid). The rest of the procedure was carried out in low-peptide binding tubes (ThermoFisher Scientific). The gel slices are destained further with 50% acetonitrile

in 25 mM ammonium bicarbonate. These gel slices are completely dried in speed-vac. These gel slices are reduced and alkylated using 10 mM dithiothreitol (DTT) and 55 mM iodoacetic acid (IAA) respectively. The gel slices dried again completely in a speed-vac (Savant SpeedVac SC 100).

For trypsin digestion, mass spectrometry / sequencing grade trypsin (Promega) is stored at a concentration of 0.1μg/μL as 12.5 μL aliquots in -80 ˚C. These aliquots are freshly thawed before use. To these aliquots, 100 μL of 25 mM ammonium bicarbonate (digestion buffer) was added to bring the final volume to 112.5 μL. The final working concentration of trypsin is 11.1 ng/μL. This trypsin digestion solution is added to the gel slices until they are completely covered and incubated on ice for 45 minutes. Any leftover solution, not absorbed by the gel-slices was removed. The digestion buffer was added to completely immerse the gel slices, and incubated overnight (14-16 hours) at 37 ˚C. A solution of 2% TFA (10 μL) was used to stop the trypsin digestion. The solution containing tryptic peptides were saved, and the gel slices containing longer peptides were sonicated in presence of 30-50 μL of 50% acetonitrile/5% formic acid followed for 15-20 minutes. This step was repeated twice for maximum peptide recovery. All the tryptic peptides were finally into a single vial and concentrated in speed-vac. Once the final volume was about 10 μL, 5 μL of 5% formic acid was added and stored in freezer until use. Sending the sample through MALDI tested the abundance and size distribution of the peptides.

2.5 Chromatography and Mass Spectrometry

After digestion, the proteolytic peptides were processed using a C18 ZipTip® (Millipore) before mass analysis using standard protocols. Reversed-phase chromatography was performed using Waters' XBridge HPLC columns (Frankel et al., 2012). Briefly, the C18 columns, 3.5-μm, 2.1 × 100 mm were used with a flow rate of 200 μL/min. The mobile phases consisted of 95:5 water : acetonitrile with 0.1% formic acid aqueous phase and a 95:5 acetonitrile : water with 0.1% formic

acid organic phase. The organic phase composition was 10% for the first 5 min, incremented to 20% for the next 10 min (total time 15 min), then to 50% for the next 25 min (total time 40 min), incremented to 80% in the next 35 min (total time 75 min), held at 80% for 10 min after which, it was again ramped to 10% in 5 min and a 10 min hold to equilibrate the column (total time 100 min).

The sample from HPLC was automatically injected into ESI inlet system. The inlet system was enclosed in a jacket known as an ESI probe. The inlet system employed a metal needle in which a capillary containing fine fused silica was inserted, in this capillary the analyte solution from HPLC was pumped in. The capillary was set at a temperature of 275 ˚C and the source voltage at 5 kV. ESI probe contaied two gases (a sheath gas and an auxiliary gas). Sheath gas runs in the center of the nozzle, but outside the needle and came in contact with this gas only at the end of the metal needle. The auxiliary gas ran in the ESI nozzle (through and outer ring). These gases were set at a pressure of 18 and 5 respectively. Nitrogen was used as both sheath (for droplet formation) and auxiliary (for humidity control) gases. The FT-ICR was set to positive ion scan at a 100K resolution (100K at m/z 400). Six largest ion scans are acquired at a time after ESI (precursor ions). The machine settings for CID (Thermo Scientific) scans were set to an isolation width of 2 and normalized collision energy of 35. Once the tandem mass spectra for these ions were acquired twice, these ions were set in exclusion list for 30 seconds. Charge state screening is performed using monoisotopic precursor selection. The tandem mass mass spectrometry was performed in collaboration with Dr. Larry Sallans and Dr. Patrick Limbach from the University of Cincinnati, Rieveschl Marshall Wilson mass spectrometry facility.

2.6 Determination of Crosslinks

Identification and analysis of peptides containing crosslinked products based on their mass modifications were performed using two softwares, (1) MassMatrix online and PC search engines (Xu and Freitas, 2007a, 2009) and (2) StavroX (Gotze et al., 2012).

A FASTA library containing PsbO, PsbP, PsbQ, and PsbR proteins was searched for both 1M NaCl and 1M CaCl2 extracts in BS3 and EDC crosslinking experiments. Both the experiments, had a decoy library that contained the same proteins but with reversed amino acid sequences. For the identification of crosslinked products using MassMatrix search engine, peptides were selected if their *P* value was ≤0.001, and for StavroX, their scores were required to be 3X of their false discovery rate (FDR). In each case, the peptides were required to exhibit 0% hits to the decoy library for further consideration.

2.7 MassMatrix Search Parameters

MassMatrix is a probabilistic scoring model which is very sensitive in providing mass accuracy. It is a statistical tool that helps in searching for disulfide and chemically crosslinked species. Additionally, it can also be used to find mass modification in case of post-translational modifications including oxidative modifications. The following settings were used for the BS3 crosslinking data analysis.

Variable modification: only methionine, Maximum number of variable PTM/peptide: 2

Mass Spectrometer Parameters: peptide tolerance 5ppm, MS/MS tolerance 0.8 Da

Crosslink Search options: BS3 (K, T, Y, S and N-terminus)

Crosslink Search mode: exploratory, Crosslink sites cleavability: Non-cleavable by enzyme

Max # of crosslink per peptide: 2, Enzymes: Trypsin, chymotrypsin, Tryp/Lys C

Fixed Modifications: Iodoacetic acid, Carboxymethyl of C

Parameters for output: Peptide length 3 to 40 amino acid

Min pp score: 5.3, Min pp_tag score: 1.3, Quantitation options: none

A FASTA library of up to three proteins was used in the PC version of MassMatrix, and any number of proteins in case of web-version of MassMatrix.

2.8 Modeling of PsbP N-terminus

We used a program MODELLER to add secondary structure to the crystallographically unresolved N-terminus of PsbP. This process required three files for the N-terminal model building process. (1) An alignment file for the template sequence and target sequence. (2) Structure file (.pdb) for the template protein. (3) A script file, that assigns the tasks to the program MODELLER. The generated models were validated for their secondary structures by Ramachandran analysis using the program PROCHECK, and analyzed further for their z-scores with the program PROSA-II. For sequence alignments, and the scripts used in protein modelling, please refer to the Appendix.

2.9 Testing Refined Models for their Stereochemistry

The models obtained for PsbP and PsbO, unconstrained and distance-constrained (as both helix and random coil at N-terminus for PsbP), were optimized to minimize their energies. The best structures were carefully selected based on their DOPE scores, which refers to discrete optimized protein energy, or the energy of the protein models, which satisfied the spatial restraints. DOPE is an atomic distance-dependent statistical potential (Shen and Sali, 2006), which calculates these scores considering the sample's native structure and atom types. It is based on a reference theory that assumes that the reference state has a uniform density and finite size. Thus these scores are dependent on the non-interacting atoms in a uniform/homogenous sphere obtained from the native structures of the proteins in a size-dependent manner.

References

Arnon, D.I. (1949). Copper enzymes in isolated chloroplasts. polyphenoloxidase in *Beta vulgaris*. Plant Physiology *24*, 1-15.

Delepelaire, P., and Chua, N.H. (1979). Lithium Dodecyl Sulfate-Polyacrylamide Gel-Electrophoresis of Thylakoid Membranes at 4-Degrees-C - Characterizations of 2 Additional Chlorophyll - Protein Complexes. Proceedings of the National Academy of Sciences of the United States of America *76*, 111-115.

Frankel, L.K., Sallans, L., Limbach, P.A., and Bricker, T.M. (2012). Identification of oxidized amino acid residues in the vicinity of the Mn(4)CaO(5) cluster of Photosystem II: implications for the identification of oxygen channels within the Photosystem. Biochemistry *51*, 6371-6377.

Gotze, M., Pettelkau, J., Schaks, S., Bosse, K., Ihling, C.H., Krauth, F., Fritzsche, R., Kuhn, U., and Sinz, A. (2012). StavroX--a software for analyzing crosslinked products in protein interaction studies. Journal of the American Society for Mass Spectrometry *23*, 76-87.

Kalkhof, S., and Sinz, A. (2008). Chances and pitfalls of chemical cross-linking with amine-reactive N-hydroxysuccinimide esters. Analytical and Bioanalytical Chemistry *392*, 305-312.

Lichtenthaler, H.K. (1987). Chlorophylls and Carotenoids - Pigments of Photosynthetic Biomembranes. Method Enzymol *148*, 350-382.

Madler, S., Bich, C., Touboul, D., and Zenobi, R. (2009). Chemical cross-linking with NHS esters: a systematic study on amino acid reactivities. Journal of Mass Spectrometry *44*, 694-706.

Shen, M.Y., and Sali, A. (2006). Statistical potential for assessment and prediction of protein structures. Protein Science : a publication of the Protein Society *15*, 2507-2524.

Sinz, A. (2006). Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions. Mass Spectrometry Reviews *25*, 663-682.

Smith, P.K., Krohn, R.I., Hermanson, G.T., Mallia, A.K., Gartner, F.H., Provenzano, M.D., Fujimoto, E.K., Goeke, N.M., Olson, B.J., and Klenk, D.C. (1985). Measurement of Protein Using Bicinchoninic Acid. Analytical Biochemistry *150*, 76-85.

Staros, J.V. (1982). Membrane-impermeant, cleavable cross-linkers: new probes of nearest neighbor relationships at one face of a membrane. Biophysical Journal *37*, 21-22.

Swaim, C.L., Smith, J.B., and Smith, D.L. (2004). Unexpected products from the reaction of the synthetic cross-linker 3,3'-dithiobis(sulfosuccinimidyl propionate), DTSSP with peptides. Journal of the American Society for Mass Spectrometry *15*, 736-749.

Xu, H., and Freitas, M.A. (2007). A mass accuracy sensitive probability based scoring algorithm for database searching of tandem mass spectrometry data. BMC Bioinformatics *8*.

Xu, H., and Freitas, M.A. (2009). MassMatrix: a database search program for rapid characterization of proteins and peptides from tandem mass spectrometry data. Proteomics *9*, 1548-1555.

CHAPTER 3

USE OF PROTEIN CROSSLINKING COUPLED TO TANDEM MASS SPECTROMETRY TO ELUCIDATE PsbP AND PsbQ INTERACTIONS IN SPINACH PHOTOSYSTEM II

3.1 Introduction

PSII is a water splitting enzyme that works as a light energy driven water-plastoquinone oxidoreductase. It consists of at least 20 proteins, of which 17 of them are intrinsic membrane proteins and at least 3 of them are very important extrinsic components required for water oxidation. In higher plants and algae, PSII subunits are encoded by both chloroplast and nuclear genomes. The intrinsic membrane proteins of PSII D1, CP47, CP43, D2, Cyt b*559* α and β are encoded by the chloroplast genes *psbA, psbB, psbC, psbD, psbE* and *psbF* respectively. However the plant nuclear genes encode the extrinsic proteins PsbO, PsbP, PsbQ and PsbR. The extrinsic proteins PsbO, PsbP and PsbQ are absolutely required for maximal rates of oxygen evolution at physiological concentrations of inorganic cofactors, Ca^{+2} and Cl^{-}. They are critically important in protecting the Mn_4CaO_5 (metal cluster) from exogenous reductants.

The cyanobacterial PSII structure has been resolved at 1.9Å resolution; and acts as a primary reference in understanding the structure and location of PSII subunits in higher plants. Due to differences in the cyanobacterial and higher plant PSII, these structures are not always appropriate for higher plant PSII studies; nevertheless, they still are very good starting point to understand the organization of higher plant PSII proteins. The intrinsic membrane core proteins are very similar in both of these species, with 80% sequence similarity. On the lumenal face however, the protein composition is different. Cyanobacteria contains PsbO, PsbU, PsbV, CyanoP, CyanoQ and Psb27 on their lumenal side of PSII while higher plants contain PsbO, PsbP and PsbQ. CyanoP and CyanoQ are structural homologs of PsbP and PsbQ. Functionally however, PsbU and PsbV from

cyanobacteria are similar to higher plant PsbP and PsbQ. Additionally, cyanobacterial X-ray crystal structure of PSII consists of PsbO, PsbU and PsbV lumenal proteins only (and not CyanoP and CyanoQ), and thus locations of other proteins remain obscure, and the binding location of higher plant extrinsic proteins is unknown. In this section, results from higher plant PsbP and PsbQ proteins are discussed.

High resolution crystal structures for isolated PsbP are available from tobacco (*Nicotiana tobaccum* PDB ID: 1V2B (Ifuku et al., 2004)), spinach (*Spinacia oleracea* PDB ID: 2VU4 (Kopecky et al., 2012)) cyanoP of cyanobacteria (structurally similar to PsbP, PDB ID: 2XB3 (Michoux et al., 2010)). Alignment of all these structures showed at least three unresolved domains. The N-terminal unresolved segment is particularly interesting, this region is suggested to interact with PSII membrane proteins by crosslinking studies and mutants analysis in spinach (cite). These structures were obtained from the soluble proteins overexpressed in *E.coli* and hence they do not represent the actual conformations in their bound state to PSII. Recently, a new crystal structure of PsbP from spinach is available at 1.8 Å resolution became available (Cao et al., 2015), this structure contains two Mn ions bound to PsbP. Two of the previously unresolved segments, ^{90}K-^{107}V and ^{134}R-^{139}D were found to be stabilized in the presence of Mn ions, and hence were resolved as loops in this structure, 4RTI. PsbP binding to PSII induces conformational changes around the Mn_4CaO_5 cluster (Tomita et al., 2009). Absence of PsbP (along with PsbQ) leads to deviation in oxygen evolution ability and hence alterations to electron transport from Q_A^- to Q_B (Roose et al., 2010; Yi et al., 2007). The N-terminus of PsbP is important for its binding to PSII and oxygen evolution activity. Two mutants, Δ19-PsbP and Δ15-PsbP could not activate oxygen evolution at physiological calcium concentrations. Additionally, it was recently determined that the binding of PsbP induces conformational changes in PsbE (α-subunit of Cyt b_{559}, a core subunit of PSII) (Nishimura et al.,

2016). EDC crosslinker studies identified a crosslink-residue pair between ^{1}A PsbP and ^{57}D PsbE residues (Ido et al., 2012). EDC is a zero length crosslinker, suggesting that these two residues are in van der Waals contact. Additionally, the functional importance of the PsbP C-terminal residues in proper positioning of N-terminus has been hypothesized to support its interaction with Cyt b_{559} (Ido et al., 2012). Release-reconstitution experiments demonstrated that binding of full length PsbQ to Δ19-PsbP could partially restore calcium retention, and oxygen evolution process (although FTIR studies have not been done on this mutant to understand structural changes around the Mn4CaO5 cluster). Removal of PsbP and PsbQ in higher plants induced a conformational change that caused CP29 to move 12 Å towards the PSII core in the super-complexes (Boekema et al., 2000). These results further support the importance of PsbP and PsbQ in structural organization of PSII.

The X-ray crystallography of the extrinsic proteins PsbP and PsbQ in their PSII-bound forms is not available, and hence the locations of these proteins remain unknown. Cyanobacterial PSII crystal structure provided enormous information about the intrinsic membrane proteins and the location of PsbU, PsbV and PsbO. This X-ray crystal structure, 3WU2, lacks CyanoP and CyanoQ proteins (the structural homologs of PsbP and PsbQ). Protein crosslinking studies in cyanobacteria (Liu et al., 2014) suggests the interaction of CyanoQ with PsbO and CP47, which may or may not bind to PSII in the same location as the higher plant to PsbP and PsbQ components. The location of PsbQ was modelled in the groove formed by each PSII monomer, such that the N-terminus of CyanoQ is located close to CP47 and PsbO, however N-terminal region of CyanoQ is different from that of higher plant PsbQ (Liu et al., 2014). This however, is not directly applicable to higher plant PsbQ, since CyanoQ and PsbQ are only structurally similar, however PsbQ is functionally more comparable to PsbV of cyanobacteria. Thus the crystal structure 3WU2 is only a good reference to understand structural organization of intrinsic membrane proteins which are about 80% similar to

their higher plant forms. Hence, binding location of higher plant extrinsic proteins to PSII is still unavailable and needs future attention.

In this dissertation, spinach is used as a model organism for higher plants to understand the relative positioning of these extrinsic proteins to PSII. The method used for this process is protein crosslinking coupled to high-resolution tandem mass spectrometry. The results from the crosslinker BS3 related to PsbP and PsbQ are shown in this section. These results have been published in the *Proceedings of the National Academy of Sciences*, USA (Mummadisetti et al., 2014). This work was performed in collaboration with Dr. Patrick Limbach and Dr. Larry Sallans, University of Cincinnati. In this dissertation, higher plant PsbP is shown in different shades of pink and magenta, PsbQ in different shades of green and PsbO in shades of blue (shown in chapter 4).

Our protein crosslinking results on the PsbP and PsbQ extrinsic proteins indicate, (1) the N-terminus of PsbP, unresolved in the current X-ray crystal structures, forms a compact structure and associates with the C-terminal domain of the protein (2) the interacting domains of PsbP and PsbQ forming a framework in PSII (3) interacting domains and orientation of the two copies of PsbQ from different PSII monomers.

Results and Discussion

3.2 BS3 CrossLinking reveals Inter- and Intra-Protein Associations

The FIGURE. 3.1 illustrate the results observed upon treatment of PS II membranes with varying concentrations (0 mM – 10 mM) of the crosslinker BS3. These membranes were washed with 1.0 M NaCl to extract PsbP, PsbQ and their crosslinked products (not crosslinked to the

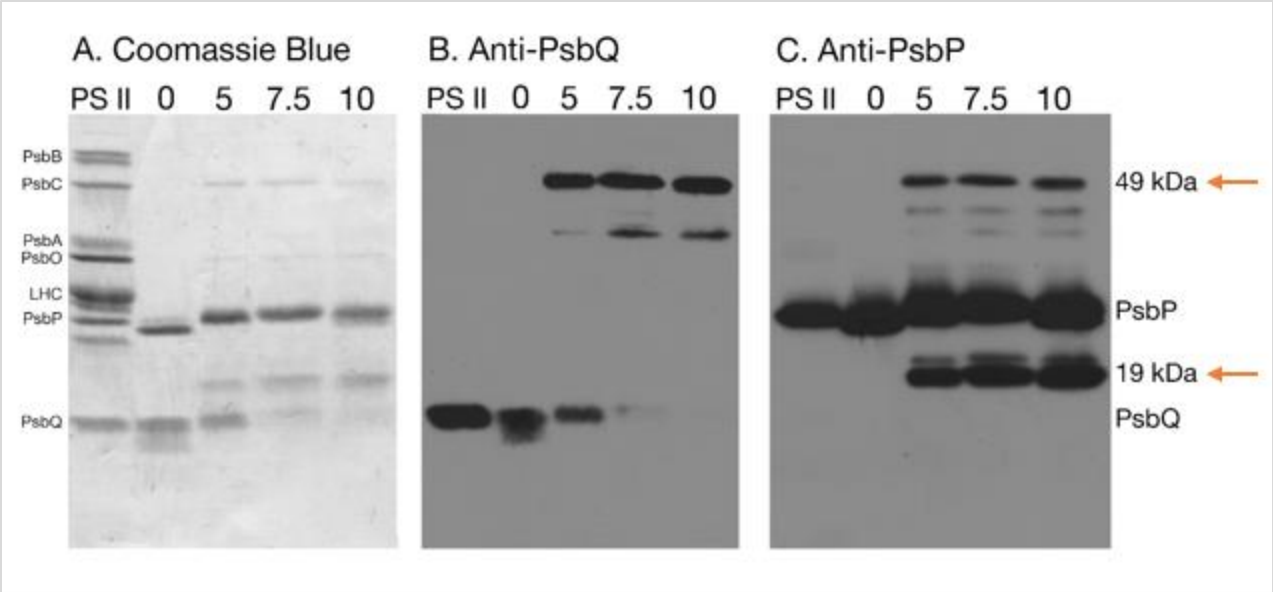


FIGURE 3.1. The BS3 Crosslinking of PSII-bound PsbP and PsbQ.

Crosslinker concentrations were 0–10 mM BS3. Illustrated are "Western" blots of PS II membranes (PS II), crosslniked and dialyzed overnight after 1 M NaCl soluble proteins obtained after crosslinking of the PS II membranes at the indicated BS3 concentrations. **A**. Overall protein profile of the samples, using coomassie blue stain. **B** and **C** identify the proteins that cross-react with anti-PsbQ and anti-PsbP, respectively. A number of PSII proteins are labeled to the left of panel *A* for reference. To the right of panel *C*, crosslinked products are labeled. Lanes were loaded with 10 μg of protein except the PS II lane, which was loaded with 10 μg of chl.

membrane proteins or PsbO). These proteins in the soluble phase were concentrated and analyzed by Lithium dodecyl sulfate (LiDS)–PAGE followed by immunoblotting. Panel A shows coomassie-stained proteins. Panel B shows bands for the proteins that reacted with anti-PsbQ and panel C is probed with anti-PsbP. As expected, in the absence of BS3, no crosslinked products were observed, and only PsbP and PsbQ were removed from the membranes by high salt (1M NaCl) treatment. Increasing concentrations of the crosslinker (5–10 mM) led to increased accumulation of a number of putative crosslinked products (19, 35, 38, and 44 kDa), which were released from the PS II membranes by salt washing. Immunoblot analysis indicated that the 19-kDa band reacted only with anti-PsbP, whereas the 35-, 38-, and 44-kDa bands reacted with both anti-PsbP and anti-PsbQ. The

crosslinked products obtained by treatment with 5 mM BS3 were chosen for further analysis. At 0 mM BS3, both PsbP and PsbQ appeared to exhibit proteolysis in the NaCl-wash extracts. It had been shown previously that both of these proteins are susceptible to proteolytic attack (Kuwabara et al., 1986; Miyao et al., 1988). The presence of the crosslinker BS3, however, appeared to suppress this endogenous proteolytic activity. The crosslinker BS3 has been used extensively in protein crosslinking studies in a variety of systems. The N-hydroxy-sulfosuccidimyl leaving group reacts with primary amines and, at low pH, with the hydroxyl groups found in tyrosyl, threonyl, and seryl residues (Sinz, 2006). The modification of threonyl and seryl hydroxyl groups is relatively inefficient and may exhibit lower stability (GT, 1996; Sinz, 2006); consequently, in this study, we have considered only crosslinks involving lysyl, tyrosyl and threonyl residues and the unblocked N termini.

The TABLE 3.1 shows the crosslinked residues identified in this study by mass spectrometry. Eleven intra-chain crosslinked products were identified within PsbP, the vast majority of these being found in the 19-kDa crosslinked protein band. Two crosslinked pairs of residues involving PsbP and PsbQ were also identified in the 44-kDa protein band. Finally, three crosslinked residue pairs of PsbQ also were identified from the 44-kDa crosslinked product. All of the crosslinked peptides identified exhibited extremely low *P* values ranging from 2.5×10^{-4} to 1.9×10^{-18}. The *P* value of $\leq 1 \times 10^{-3}$ is used as a criteria for the crosslinked residues in this study. The quality of the mass spectrometry is illustrated in the FIGURE 3.2, with a median *P* value of from the results of crosslinked residue with the median *P* value of 7.9×10^{-11}. Additionally, the data for highest *P* value (2.5×10^{-4}) is shown in the FIGURE 3.3 and the lowest *P* value 1.9×10^{-18} in the FIGURE 3.4.

TABLE 3.1. Crosslinked Residues of the PSII-Bound PsbP and PsbQ Proteins.

The table consists of all the crosslinks that qualified the criteria of *p value* $\leq$ 0.001. Crosslinks within PsbP forming intramolecular crosslinks are shown as PsbP-PsbP crosslinks, intermolecular crosslinks are shown as PsbP-PsbQ crosslinks. Internal PsbQ crosslinks consist of crosslinks with both intra- and inter-molecular crosslinks. The values are mapped back to their respective crystal structures.

	Cross-linked Residues	[1]P-Score	[2]Crystal Structure Distance
Internal PsbP cross-links	^{1}A-^{174}K	7.0 x 10^{-5}	NA
	^{1}A-^{170}K	3.9 x 10^{-13}	NA
	^{1}A-^{173}K	1.5 x 10^{-11}	NA
	^{2}Y-^{170}K	7.9 x 10^{-11}	NA
	^{11}K-^{13}K	1.2 x 10^{-13}	NA
	^{11}K-^{14}K	1.6 x 10^{-8}	NA
	^{11}K-^{174}K	2.5 x 10^{-4}	NA
	^{13}K-^{174}K	6.3 x 10^{-5}	NA
	^{14}K -^{174}K	5.0 x 10^{-16}	NA
	^{33}K -^{174}K	6.3 x 10^{-9}	9.3 Å
	^{40}K-^{155}K	1.9 x 10^{-10}	14.4 Å
	^{40}K-^{157}K	1.2 x 10^{-9}	16.9 Å
PsbP – PsbQ cross-links	PsbP:^{93}Y-PsbQ:^{1}E	2.0 x 10^{-13}	NA
	PsbP:^{96}K-PsbQ:^{1}E	1.9 x 10^{-18}	NA
	PsbP:^{97}T-PsbQ:^{1}E	6.3 x 10^{-18}	NA
Internal PsbQ cross-links	^{53}K-^{96}K	2.0 x 10^{-13}	8.5 Å
	^{98}K-^{133}Y	1.0 x 10^{-8}	33.1 Å
	^{101}K-^{133}Y	1.0 x 10^{-15}	30.0 Å
	^{125}K-^{133}Y	5.0 x 10^{-15}	16.9 Å
	^{128}T-^{133}Y	7.9 x 10^{-15}	9.0 Å
	^{126}S -^{133}Y	3.16 x 10^{-14}	8.0 Å

NA, not applicable, as residues were not resolved in the crystal structure.
[1]Probability that the peptide match is a random occurrence.
[2]Distances reported for the crystal structure of isolated PsbP (PDB ID code 2VU4) and PsbQ (PBD ID code 1VYK) are measured between the putative crosslinked atoms (amine nitrogens or hydroxyl oxygens).

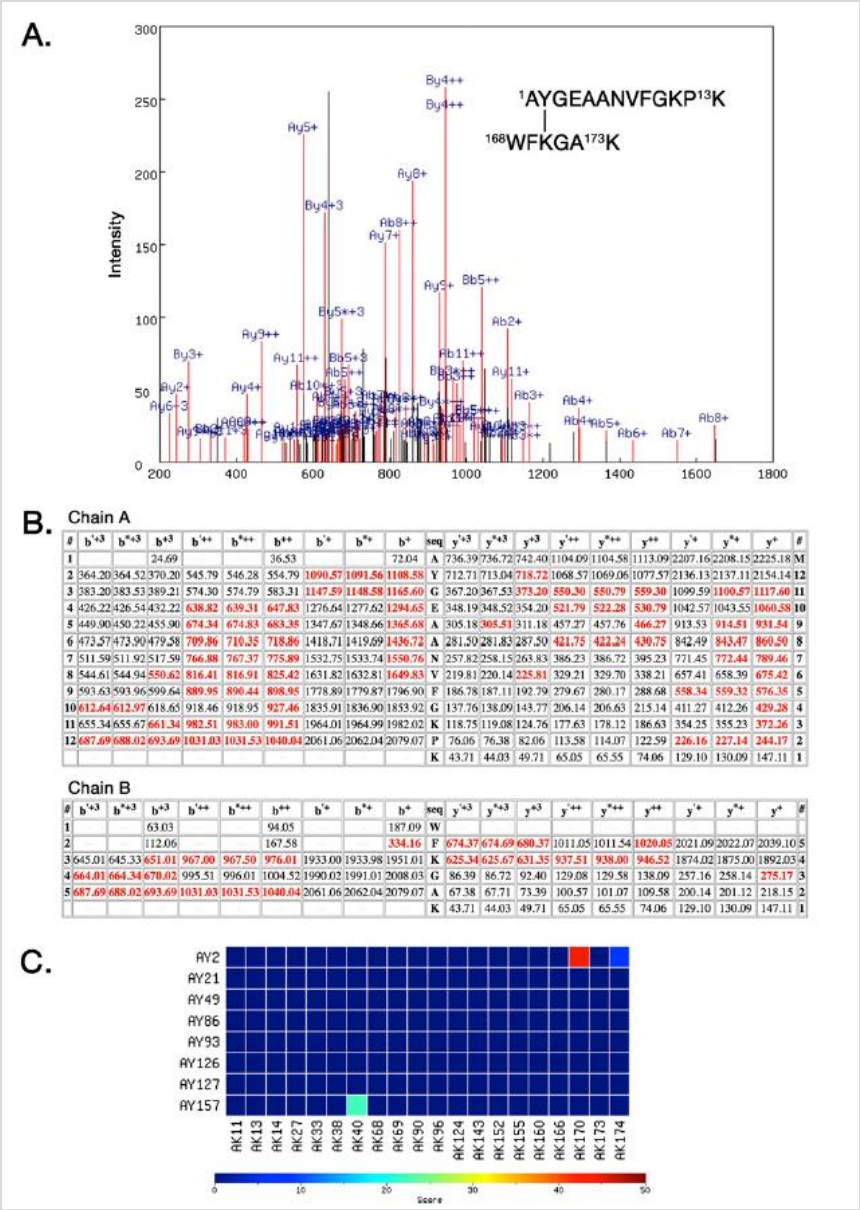


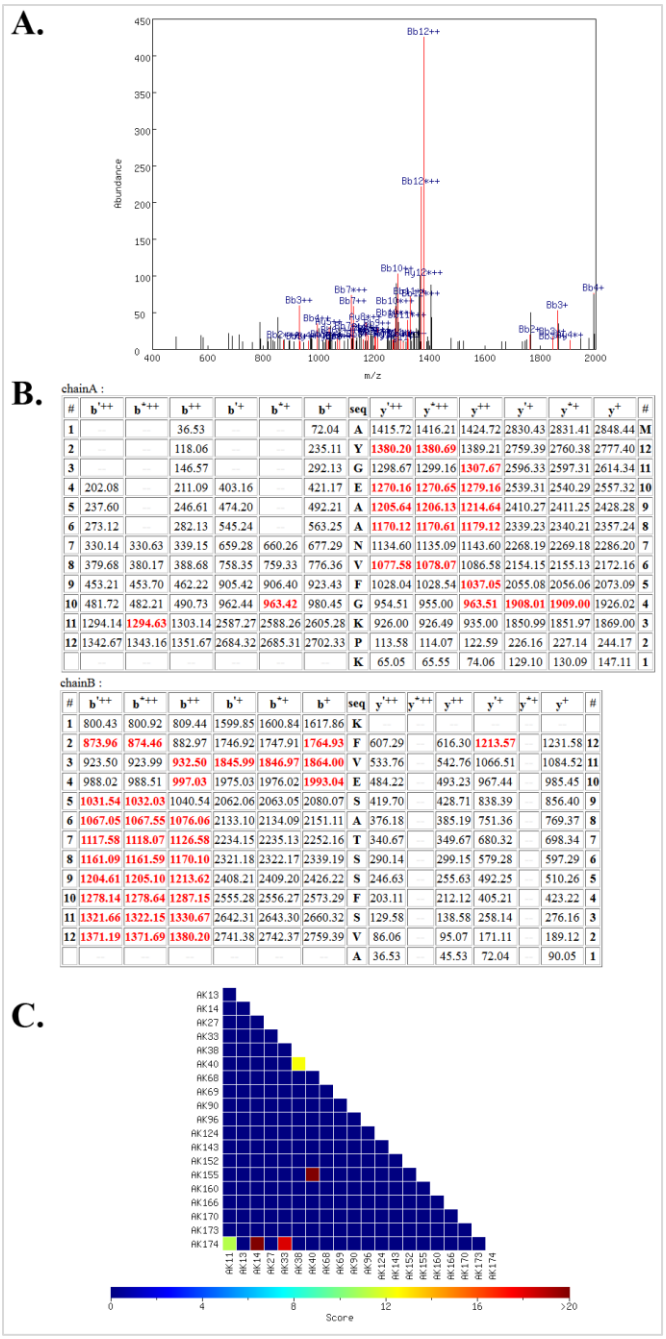
FIGURE 3.2. Crosslinked Residues with Median *P* value (7.9×10^{-11}).

Quality of the mass spectrometry used in this study. Shown are the mass spectrometry data obtained for the median *P* value of crosslinked residue pair ^{2}Y–^{170}K (TABLE 3.1; *P* value = 7.9×10^{-11}).

A. Mass spectra obtained for this peptide. Identified ions are shown in red; those not identified (predicted ions) are in black.

B. Table of predicted ions from this peptide. Identified ions are shown in red; those not identified are in black. Note the nearly complete y- and b-ion series obtained.

C. Heat map that assists in the identification of crosslinked species. The identified peptide is shown in red. The PsbP tyrosyl-residues are shown on the y axis, and lysyl residues are shown on the x axis.



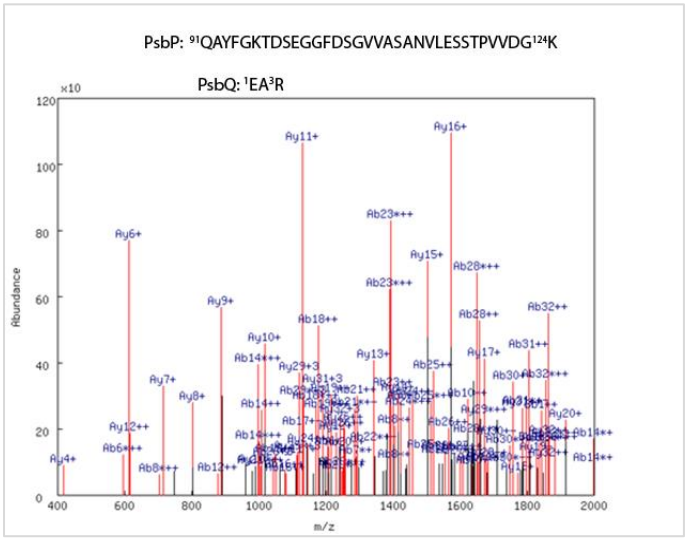
#	b^{+}	b^{++}	b^{+++}	b^{++}	b^{+}	b^{++}	b^{+}	seq	y^{+}	y^{++}	y^{+++}	y^{++}	y^{+}	y^{++}	y^{+}	#
1				36.53			72.04	A	1415.72	1416.21	1424.72	2830.43	2831.41	2848.44		M
2				118.06			235.11	Y	**1380.20**	**1380.69**	1389.21	2759.39	2760.38	2777.40		12
3				146.57			292.13	G	1298.67	1299.16	**1307.67**	2596.33	2597.31	2614.34		11
4	202.08			211.09	403.16		421.17	E	**1270.16**	**1270.65**	**1279.16**	2539.31	2540.29	2557.32		10
5	237.60			246.61	474.20		492.21	A	**1205.64**	**1206.13**	**1214.64**	2410.27	2411.25	2428.28		9
6	273.12			282.13	545.24		563.25	A	**1170.12**	**1170.61**	**1179.12**	2339.23	2340.21	2357.24		8
7	330.14	330.63		339.15	659.28	660.26	677.29	N	1134.60	1135.09	1143.60	2268.19	2269.18	2286.20		7
8	379.68	380.17		388.68	758.35	759.33	776.36	V	**1077.58**	**1078.07**	1086.58	2154.15	2155.13	2172.16		6
9	453.21	453.70		462.22	905.42	906.40	923.43	F	1028.04	1028.54	**1037.05**	2055.08	2056.06	2073.09		5
10	481.72	482.21		490.73	962.44	**963.42**	980.45	G	954.51	955.00	**963.51**	**1908.01**	**1909.00**	1926.02		4
11	1294.14	**1294.63**		1303.14	2587.27	2588.26	2605.28	K	926.00	926.49	935.00	1850.99	1851.97	1869.00		3
12	1342.67	1343.16		1351.67	2684.32	2685.31	2702.33	P	113.58	114.07	122.59	226.16	227.14	244.17		2
								K	65.05	65.55	74.06	129.10	130.09	147.11		1

#	b^{+}	b^{++}	b^{+++}	b^{++}	b^{+}	b^{++}	b^{+}	seq	y^{+}	y^{++}	y^{+++}	y^{++}	y^{+}	y^{++}	y^{+}	#
1	800.43	800.92		809.44	1599.85	1600.84	1617.86	K								
2	**873.96**	**874.46**		882.97	1746.92	1747.91	**1764.93**	F	607.29			616.30	**1213.57**		1231.58	12
3	923.50	923.99		**932.50**	**1845.99**	**1846.97**	**1864.00**	V	533.76			542.76	1066.51		1084.52	11
4	988.02	988.51		**997.03**	1975.03	1976.02	**1993.04**	E	484.22			493.23	967.44		985.45	10
5	**1031.54**	**1032.03**		1040.54	2062.06	2063.05	2080.07	S	419.70			428.71	838.39		856.40	9
6	**1067.05**	**1067.55**		**1076.06**	2133.10	2134.09	2151.11	A	376.18			385.19	751.36		769.37	8
7	**1117.58**	**1118.07**		**1126.58**	2234.15	2235.13	2252.16	T	340.67			349.67	680.32		698.34	7
8	**1161.09**	**1161.59**		**1170.10**	2321.18	2322.17	2339.19	S	290.14			299.15	579.28		597.29	6
9	**1204.61**	**1205.10**		**1213.62**	2408.21	2409.20	2426.22	S	246.63			255.63	492.25		510.26	5
10	**1278.14**	**1278.64**		**1287.15**	2555.28	2556.27	2573.29	F	203.11			212.12	405.21		423.22	4
11	**1321.66**	**1322.15**		**1330.67**	2642.31	2643.30	2660.32	S	129.58			138.58	258.14		276.16	3
12	**1371.19**	**1371.69**		**1380.20**	2741.38	2742.37	2759.39	V	86.06			95.07	171.11		189.12	2
								A	36.53			45.53	72.04		90.05	1

FIGURE 3.3 The Crosslinked Residues with the Highest *P* Value (7×10^{-4}).

Shown are the mass spectrometry data obtained for the highest *P* value (used in this study) of crosslinked residue pair from PsbP ^{11}Y–^{174}K (TABLE 3.1; *P* value = 7.9×10^{-11}). (A) Heat map assisting in the identification of crosslinked residues. (B) Mass spectra obtained for this peptide. Identified ions are shown in red; those not identified (predicted ions) are in black. (C) Table of predicted ions from this peptide. Identified ions are shown in red; those not identified are in black. Note the nearly complete y- or b-ion series obtained. The PsbP: ^{11}K is shown on the y axis, and ^{174}K on the x axis.

A



B

chainA :

#	b'+3	b"+3	b+3	b'++	b"++	b++	b'+	b"+	b+	seq	y'+3	y"+3	y+3	y'++	y"++	y++	y'+	y"+	y+	#
1		38.02	43.69		56.52	65.04		112.04	129.07	Q	1304.96	1305.29	1310.97	1956.94	1957.43	1965.95	3912.88	3913.86	3930.89	M
2		61.70	67.37		92.04	100.56		183.08	200.10	A	1262.28	1262.61	1268.28	1892.91	1893.41	1901.92	3784.82	3785.80	3802.83	33
3		116.05	121.73		173.57	182.09		346.14	363.17	Y	1238.60	1238.93	**1244.60**	1857.39	1857.89	**1866.40**	3713.78	3714.77	3731.79	32
4		165.07	170.75		247.11	255.62		493.21	510.23	F	1184.24	1184.57	**1190.25**	1775.86	1776.35	**1784.87**	3550.72	3551.70	3568.73	31
5		184.08	189.76		275.62	284.13		550.23	567.26	G	1135.22	1135.55	1141.23	1702.33	1702.82	**1711.33**	3403.65	3404.63	3421.66	30
6	397.20	397.53	403.21	595.30	**595.80**	604.31	1189.60	**1190.58**	**1207.61**	K	**1116.21**	**1116.54**	**1122.22**	**1673.82**	**1674.31**	**1682.82**	3346.63	3347.61	3364.64	29
7	430.89	431.22	436.89	645.83	646.32	654.83	1290.65	**1291.63**	1308.66	T	902.76	903.09	908.77	1353.64	1354.13	1362.65	2706.27	2707.26	2724.28	28
8	469.23	469.56	475.23	**703.34**	**703.83**	712.35	**1405.67**	**1406.66**	**1423.69**	D	869.08	869.41	875.08	1303.12	1303.61	1312.12	2605.23	2606.21	2623.24	27
9	498.24	498.57	504.24	746.86	747.35	755.86	1492.71	1493.69	1510.72	S	830.74	831.07	836.74	1245.60	1246.10	**1254.61**	2490.20	2491.18	2508.21	26
10	541.25	541.58	547.26	811.38	811.87	820.38	1621.75	**1622.73**	**1639.76**	E	801.73	**802.06**	807.73	1202.09	1202.58	1211.09	2403.17	2404.15	2421.18	25
11	560.26	560.59	566.27	839.89	840.38	848.89	1678.77	**1679.75**	1696.78	G	758.71	759.04	764.72	1137.57	1138.06	**1146.57**	2274.12	2275.11	2292.14	24
12	579.27	579.60	585.27	868.40	868.89	**877.41**	1735.79	1736.78	1753.80	G	739.71	740.03	745.71	1109.06	1109.55	1118.06	2217.10	2218.09	2235.11	23
13	628.29	628.62	634.30	941.93	942.43	950.94	1882.86	**1883.84**	1900.87	F	720.70	721.03	726.70	**1080.54**	**1081.04**	1089.55	2160.08	2161.07	2178.09	22
14	666.63	666.96	672.64	**999.45**	**999.94**	**1008.45**	**1997.89**	**1998.87**	2015.90	D	671.68	672.00	677.68	**1007.01**	1007.50	1016.02	2013.01	2014.00	2031.02	21
15	695.64	695.97	701.65	1042.96	**1043.46**	**1051.97**	2084.92	2085.90	2102.93	S	653.33	653.66	659.34	949.50	949.99	958.50	1897.99	1898.97	**1916.00**	20
16	714.65	**714.98**	720.66	1071.47	1071.97	**1080.48**	2141.94	2142.93	2159.95	G	604.32	604.65	610.33	905.98	906.47	914.99	1810.95	1811.94	**1828.97**	19
17	747.67	748.00	753.68	**1121.01**	**1121.50**	**1130.01**	2241.01	2241.99	2259.02	V	585.32	585.64	591.32	**877.47**	**877.96**	886.48	1753.93	1754.92	**1771.94**	18
18	780.70	781.03	786.70	**1170.54**	**1171.03**	**1179.55**	2340.08	2341.06	2358.09	V	552.29	552.62	558.30	827.94	828.43	836.94	**1654.86**	1655.85	**1672.88**	17
19	804.38	804.70	810.38	**1206.06**	**1206.55**	**1215.07**	2411.11	2412.10	2429.13	A	519.27	519.60	525.27	778.40	778.89	787.41	**1555.80**	**1556.78**	**1573.81**	16
20	833.39	833.72	839.39	**1249.58**	**1250.07**	**1258.58**	2498.15	2499.13	2516.16	S	495.59	495.92	501.59	742.88	743.38	751.89	1484.76	1485.74	**1502.77**	15
21	857.07	857.39	863.07	**1285.10**	**1285.59**	**1294.10**	2569.18	2570.17	2587.19	A	466.58	466.91	472.58	699.37	699.86	708.37	1397.73	1398.71	**1415.74**	14
22	895.08	895.41	901.08	1342.12	**1342.61**	1351.12	2683.23	2684.21	2701.24	N	442.90	443.23	448.91	663.85	664.34	672.85	1326.69	1327.67	**1344.70**	13
23	928.10	928.43	934.11	**1391.65**	**1392.14**	**1400.66**	2782.30	2783.28	2800.31	V	404.89	405.22	410.89	606.83	607.32	**615.83**	1212.65	1213.63	**1230.66**	12
24	965.80	966.13	971.80	**1448.19**	**1448.69**	**1457.20**	2895.38	2896.36	2913.39	L	371.86	372.19	377.87	557.29	557.78	566.30	**1113.58**	1114.56	**1131.59**	11
25	**1008.81**	**1009.14**	1014.82	**1512.71**	**1513.21**	**1521.72**	3024.42	3025.41	3042.43	E	334.17	334.50	340.17	500.75	501.24	509.76	**1000.49**	**1001.48**	**1018.51**	10
26	1037.82	1038.15	**1043.83**	**1556.23**	**1556.72**	**1565.24**	3111.45	3112.44	3129.46	S	291.16	291.48	297.16	436.23	436.72	445.23	871.45	872.44	**889.46**	9
27	1066.83	1067.16	1072.84	**1599.75**	**1600.24**	**1608.75**	3198.49	3199.47	3216.50	S	262.14	262.47	268.15	392.71	393.21	401.72	784.42	785.40	**802.43**	8
28	1100.52	1100.84	1106.52	**1650.27**	**1650.76**	**1659.28**	3299.53	3300.52	3317.54	T	233.13	233.46	239.14	349.20	349.69	358.20	697.39	698.37	**715.40**	7
29	**1132.87**	**1133.20**	1138.87	1698.80	1699.29	1707.80	3396.59	3397.57	3414.60	P	199.45	199.78	205.46	298.67	299.17	307.68	**596.34**	597.32	**614.35**	6
30	1165.89	1166.22	1171.89	**1748.33**	**1748.82**	**1757.34**	3495.66	3496.64	3513.67	V	167.10	167.43	173.10	250.15	250.64	259.15	499.29	500.27	517.30	5
31	1198.91	1199.24	1204.92	**1797.87**	**1798.36**	**1806.87**	3594.72	3595.71	3612.73	V	134.08	134.41	140.08	200.61	201.11	209.62	400.22	401.20	**418.23**	4
32	1237.25	1237.58	1243.26	**1855.38**	**1855.87**	**1864.38**	3709.75	3710.73	3727.76	D	101.06	101.38	107.06	151.08	151.57	160.08	301.15	302.13	319.16	3
33	1256.26	1256.59	1262.27	**1883.89**	**1884.38**	1892.89	3766.77	3767.76	3784.78	G	62.71	63.04	68.72	93.57	94.06	102.57	186.12	187.11	204.13	2
										K	43.71	44.03	49.71	65.05	65.55	74.06	129.10	130.09	147.11	1

chainB :

#	b'+3	b"+3	b+3	b'++	b"++	b++	b'+	b"+	b+	seq	y'+3	y"+3	y+3	y'++	y"++	y++	y'+	y"+	y+	#
1	1223.25	1223.58	1229.25	1834.37	**1834.86**	**1843.37**	3667.73	3668.71	3685.74	E										
2	1246.93	1247.25	1252.93	1869.89	1870.38	1878.89	3738.77	3739.75	3756.78	A	76.72	77.05	82.72	114.58	115.07	123.58	228.15	229.13	246.16	2
										R	53.04	53.37	59.04	79.06	79.55	88.06	157.11	158.09	175.12	1

C

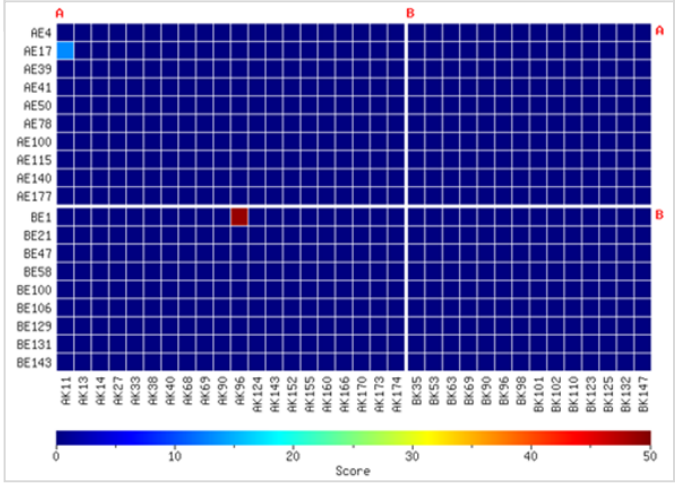


FIGURE 3.4. Crosslinked Residues with the Lowest *P* Value (1.9×10^{-18}).

Shown are the mass spectrometry data obtained for the highest *P* value (used in this study) of crosslinked residue pair from PsbP: ^{96}K– PsbQ: ^{1}E (TABLE 3.1; *P* value = 1.9×10^{-18}). (A) Mass spectra obtained for this peptide. Identified ions are shown in red; those not identified (predicted ions) are in black. (B) Heat map assisting in the identification of crosslinked residues. (C) Table of predicted ions from this peptide. Identified ions are shown in red; those not identified are in black. Note the nearly complete y- or b-ion series obtained. The PsbP: ^{11}K is shown on the y axis, and ^{174}K on the x axis.

The "Western" blots suggest that the 19-kDa band was the result of intra-chain protein crosslinking. Subsequent mass spectral analysis indicated that no proteolysis was evident, as we obtained complete MS sequence coverage for both the N- and C-termini. The three higher mass bands contained PsbP and PsbQ and represent inter-chain crosslinked products. The 19- and 44-kDa crosslinked products were most abundant and were chosen for further analysis.

3.3 The N-Terminus of PsbP Associates with C-Terminal Domains, Analysis of 19kDa Band

Elucidation of the structure of the N- terminus of PsbP is critical for understanding the function of this component. The N terminus is required for the efficient binding of PsbP to the photosystem and for its function (Ifuku and Sato, 2001) in lowering the calcium and chloride requirement for oxygen evolution (Demetrios F. Ghanotakis, 1985; Ghanotakis et al., 1984b; Miyao and Murata, 1984). Unfortunately, this important domain is not resolved in the current crystal structure of unbound PsbP (Kopecky et al., 2012) (PDB ID code 2VU4). Our observation that after crosslinking, PsbP, which normally migrates at 24 kDa, is observed to migrate at 19 kDa (FIGURE 3.1) indicates that BS3 treatment prevents complete unfolding of the protein in the presence of LiDS and suggests that PsbP, when associated with the membrane, has a compact structure. Mass spectrometry data demonstrate that the N-terminal domain (^{1}A–^{15}N) of PsbP is closely associated with the C-terminus (^{170}K–^{186}A) of the protein. Nine independent crosslinked products were identified that demonstrate this association; these are shown in TABLE 3.1. The active radius of BS3 is 11.4 Å, and hence the crosslinked residue-pairs indicate that these residues of PsbP are within 11.4 Å of each other in the PSII-bound state. This information places strong constraints on the structures of PsbP that can it can attain in its bound state. Additionally, two other residue pairs of PsbP, ^{40}K - ^{155}K and ^{40}K - ^{157}Y, were also identified. These residues are 14.3 – 16.3 Å (including rotamers) apart in the PSII unbound-PsbP (PDB: 4RTI) indicating that a conformational change of

3–5 Å occurs in this region of the protein upon its binding to PS II complex. The residue-pairs PsbP: ^{40}K – PsbP: ^{155}K and ^{40}K – PsbP: ^{155}K are showin in FIGURE 3.5 and FIGURE 3.6 respectively. Similar apparent conformational change (Ido et al., 2014) was observed between the residues ^{115}E and ^{173}K of PsbP. These residues were crosslinked with EDC, a zero-length crosslinker, indicating that a conformational change of 7.4 Å occurs in this region of the protein when it associates with the photosystem.

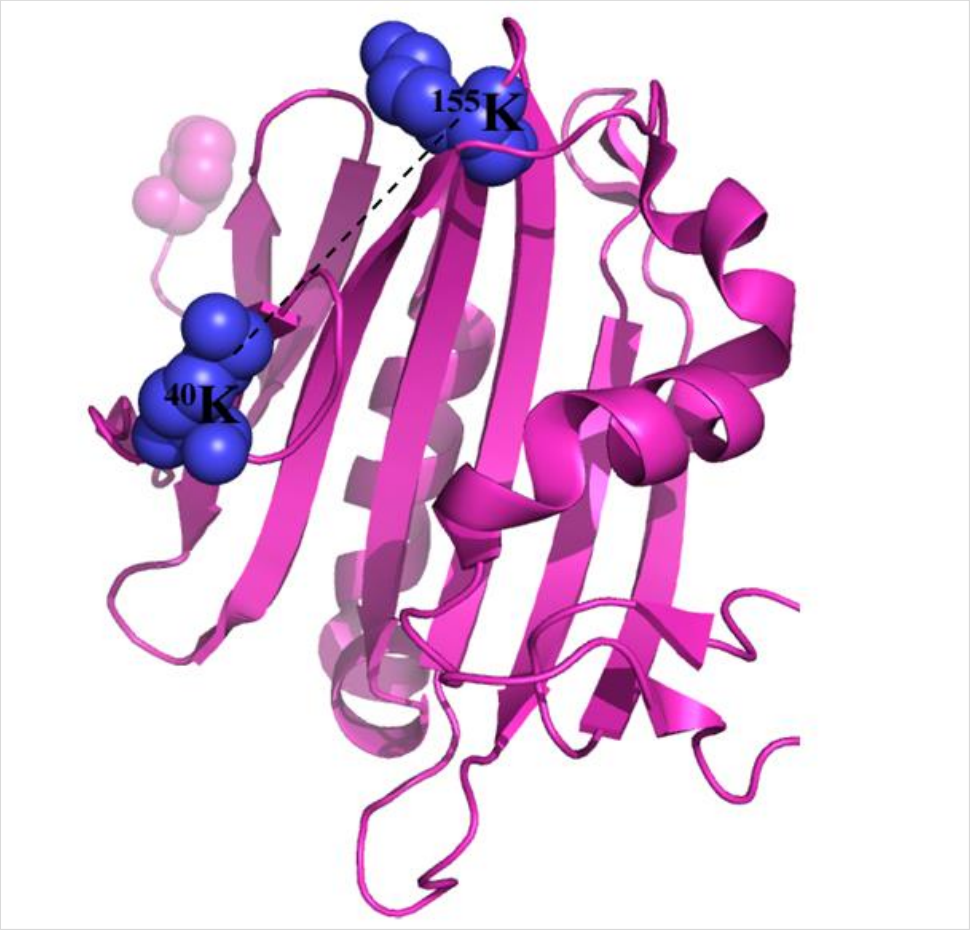


FIGURE 3.5. Conformation Changes upon PsbP Binding to PSII.

Crosslink residue-pairs ^{40}K - ^{155}K are shown in spheres. The distance between ^{40}K and ^{155}K when mapped on to the crystal structure, 4RTI, is 17 Å, and 14 Å -17 Å considering all rotamers. An increase of 2.6 Å – 6 Å distance between these residues on PsbP binding to PSII. PsbP (4RTI) structure is shown in pink, and the N-terminal residues are shown as spheres.

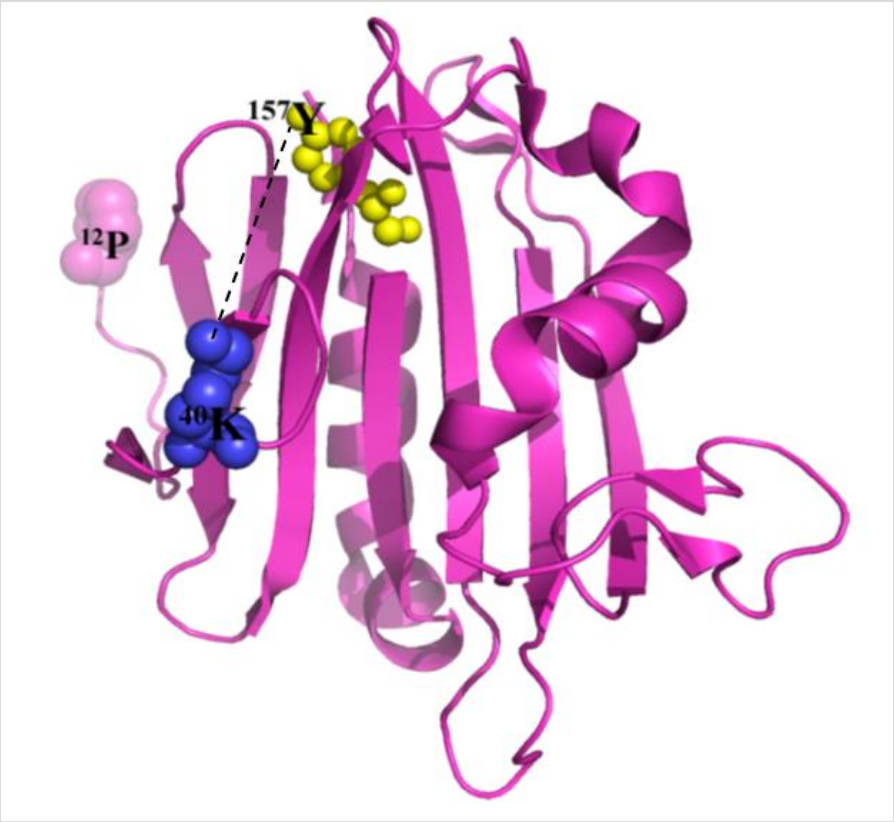


FIGURE 3.6. Conformation Changes upon PsbP Binding to PSII.

Crosslink residue pair ^{40}K - ^{155}K are shown in spheres. The distance between ^{40}K and ^{155}K when mapped on to the crystal structure, 4RTI, is 17 Å, and 14 Å -17 Å considering all rotamers. An increase of 2.6 Å – 6 Å distance between these residues on PsbP binding to PSII. PsbP (4RTI) structure is shown in pink, and the N-terminal residues are shown as spheres.

Secondary structure analysis using the Genesilico Metaserver (Kurowski and Bujnicki, 2003) of the unresolved N-terminal domain of PsbP indicates that it is likely that this region assumes either an α-helical or a random coil secondary structure. Given the strong distance constraints provided by the observed inter-chain PsbP crosslinked products, molecular dynamic refinement can provide useful models for the structure of the N-terminus of PsbP. The program MODELLER (Eswar et al., 2007) was used to produce the structures shown in FIGURES 3.7 and 3.8.

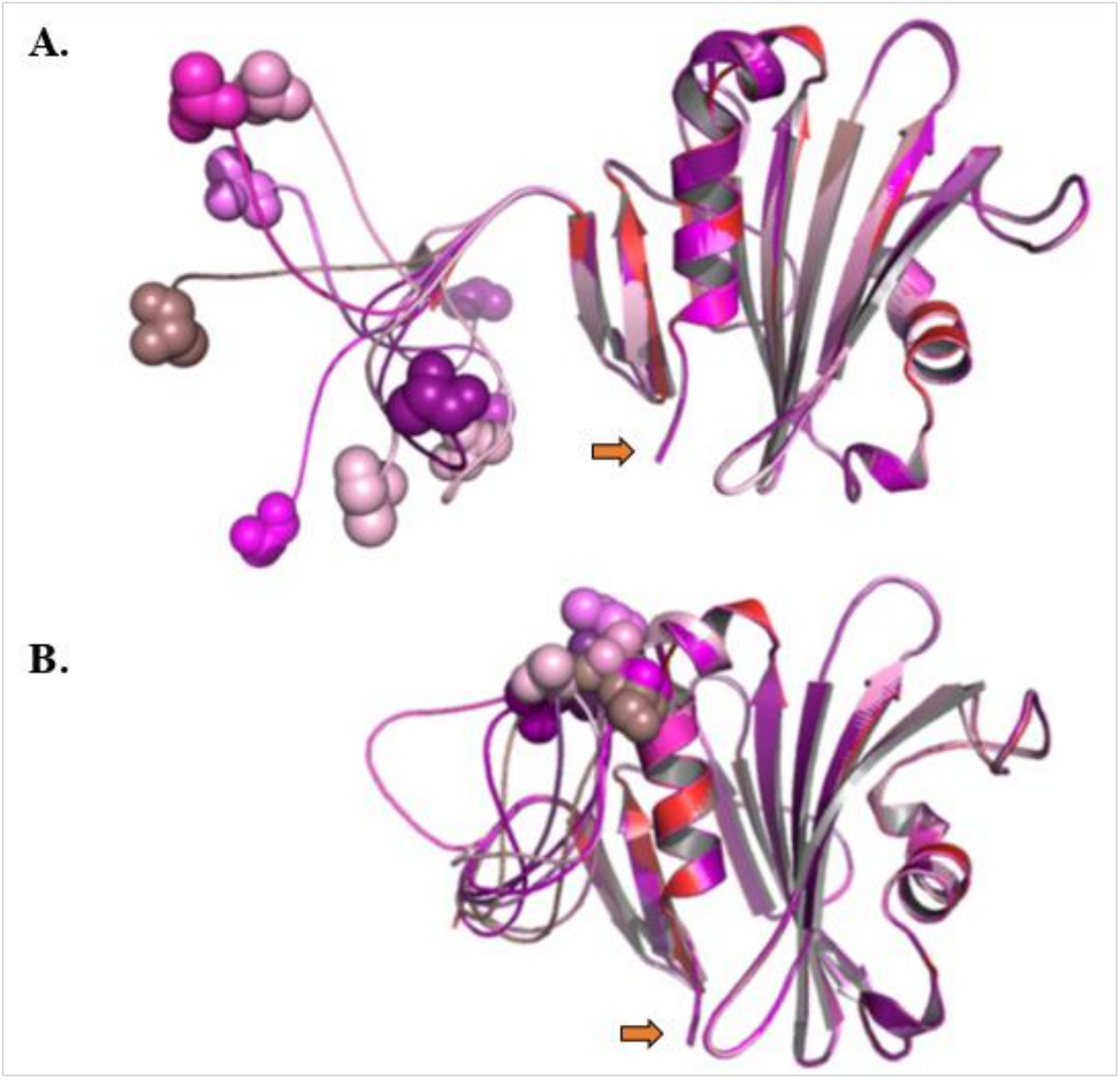


FIGURE 3.7. Molecular dynamic Refinement for the N-terminus of PsbP.

Shown are the top 10 low energy models (in different shades of purple) for both **A**. N-terminal random-coil structure shown without distance constraints, **B.** random-coil structure incorporating the distance constraints shown in TABLE 3.1. Residue ^{1}A is shown as spheres, C-terminal ^{186}A shown with orange arrows. Spinach PsbP PDB: 4RTI is shown in red.

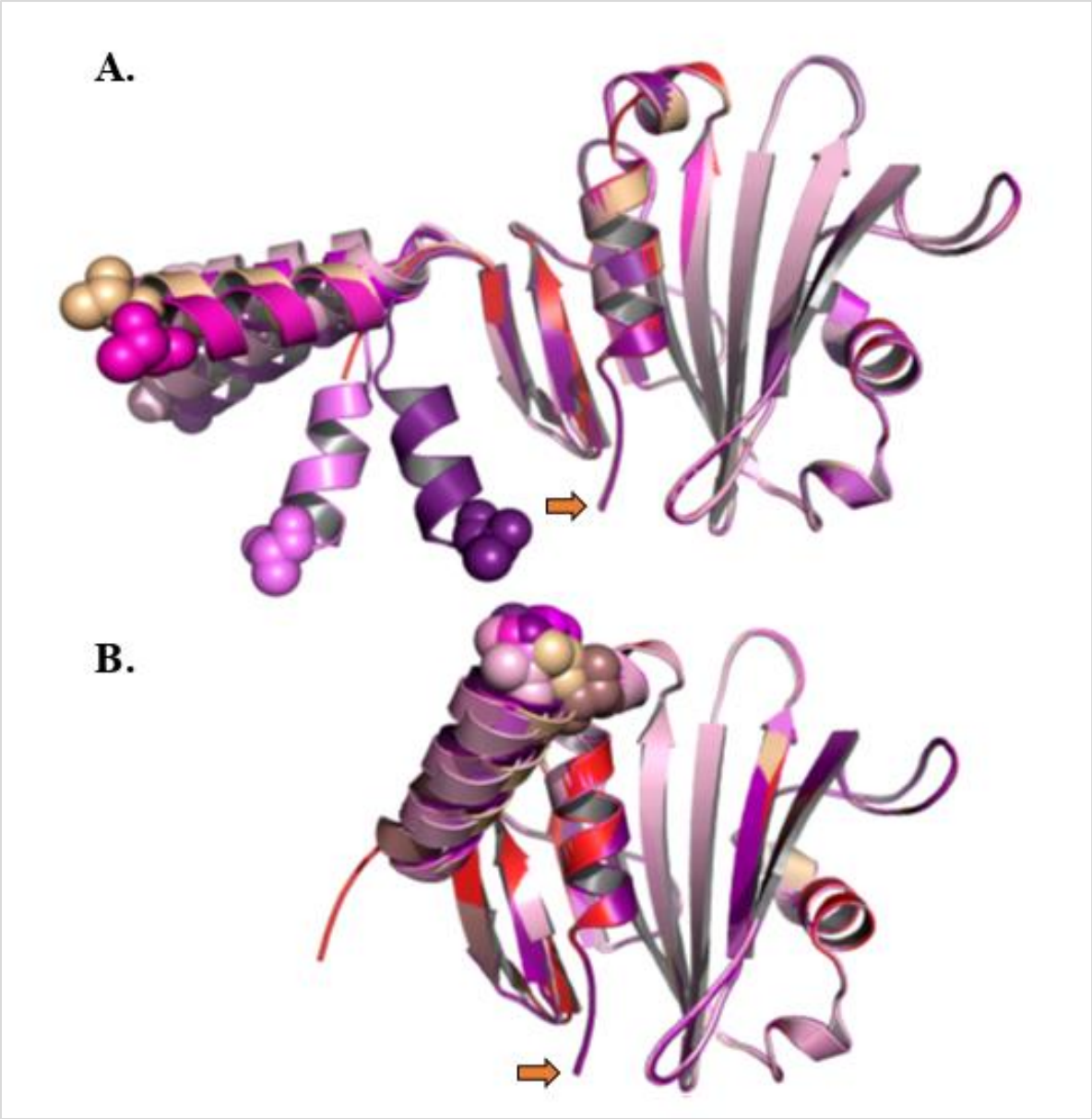


FIGURE 3.8. Molecular Dynamic Refinement of the N-terminus of PsbP

Shown are the top 10 low energy models (shown in different shades of purple) for both **A**. N-terminal α-helices shown without distance constraints, **B.** α-helices incorporating the distance constraints shown in TABLE 3.1. Residue ^{1}A is shown as spheres, C-terminal ^{186}A shown with orange arrows. Spinach PsbP PDB: 4RTI is shown in red.
Similar to FIGURE 3.7 (B), ^{1}A is located in very similar positions, again indicating that the N terminus of PSII-bound PsbP has a compact structure.

In both instances, all 10 predicted structures exhibit similar low discrete optimized protein energy scores (∼–17,000). Importantly, regardless of the type of N- terminus modeled (α-helical or random coil), the location of the N-terminal residue ^{1}A is located in very similar positions. This data indicates that PsbP assumes a compact structure when the protein is associated with the PS II complex. It had been suggested that the unresolved N- terminus might form an extended structure when associated with the PS II core complex (Bricker et al., 2012; Ido et al., 2014); this, apparently, is not the case. Residue ^{1}A has been shown to directly interact with PsbE: ^{57}E (Ido et al., 2012; Ido et al., 2014); consequently, our results indicate that PsbP must be in close proximity to PsbE and CP47.

The active radius of BS3 is 11.4 Å, and hence the crosslinked residue-pairs indicate that these residues of PsbP are within 11.4 Å of each other in the PSII-bound state. This information places strong constraints on the structures of PsbP that can it can attain in its bound state. Additionally, two other residue pairs of PsbP, ^{40}K - ^{155}K and ^{40}K - ^{157}Y, were also identified. These residues are 14.3 – 16.3 Å (including rotamers) apart in the PSII unbound-PsbP (PDB: 4RTI) indicating that a conformational change of 3–5 Å occurs in this region of the protein upon its binding to PS II complex. The residue-pairs PsbP: ^{40}K – PsbP: ^{155}K and ^{40}K – PsbP: ^{155}K are showin in FIGURE 3.5 and FIGURE 3.6 respectively. Similar apparent conformational change (Ido et al., 2014) was observed between the residues ^{115}E and ^{173}K of PsbP. These residues were crosslinked with EDC, a zero-length crosslinker, indicating that a conformational change of 7.4 Å occurs in this region of the protein when it associates with the photosystem.

3.4 PsbP and PsbQ Are Closely Associated

Both PsbP and PsbQ bind to PS II with extremely high affinity (low nM K_ds), and their removal requires high salt concentrations or markedly elevated pH conditions (Mitsue Miyao, 1989). The direct interaction between PsbP and PsbQ has been proposed by a number of investigators. Reconstitution studies have indicated that PsbQ cannot bind to the photosystem in the absence of PsbP (Ghanotakis, 1991; Mitsue Miyao, 1989). PsbQ appears to stabilize the structural and functional interaction of PsbP to PS II. An N-terminal 15-amino-acid residue-truncated PsbP cannot bind to PS II in the absence of PsbQ but can bind in its presence. And most importantly, in the presence of PsbQ, the function of the truncated PsbP is partially restored (Ifuku and Sato, 2002; Kakiuchi et al., 2012) suggesting that PsbQ probably helps in aligning and proper binding of PsbP to PSII.

Analysis of the 44-kDa crosslinked band (FIGURE 3.1) allowed the identification of an interacting domain between the PsbP and PsbQ subunits. Three residues PsbP:^{97}T, PsbP:^{96}K and PsbP:^{93}Y were crosslinked to PsbQ:^{1}E within this domain with the BS3 crosslinker, indicating that these residues of PsbP and PsbQ are within 11.4 Å (TABLE 3.1). The FIGURE 3.9 illustrates one of the many possible models of this interaction. These residues of PsbP, ^{93}Y, ^{96}K, and ^{97}T are located in the 17-residue loop 3A of PsbP (^{89}G–^{105}S), which is found crosslinked to the N-terminal ^{1}E of PsbQ. This loop of PsbP was unresolved in the previous crystal structure of the unbound PsbP (PDB: 2VU4) however, was apparently stabilized in presence of manganese and, consequently, the structure of this loop domain has been determined in the current PsbP crystal structure (PDB: 4RTI).

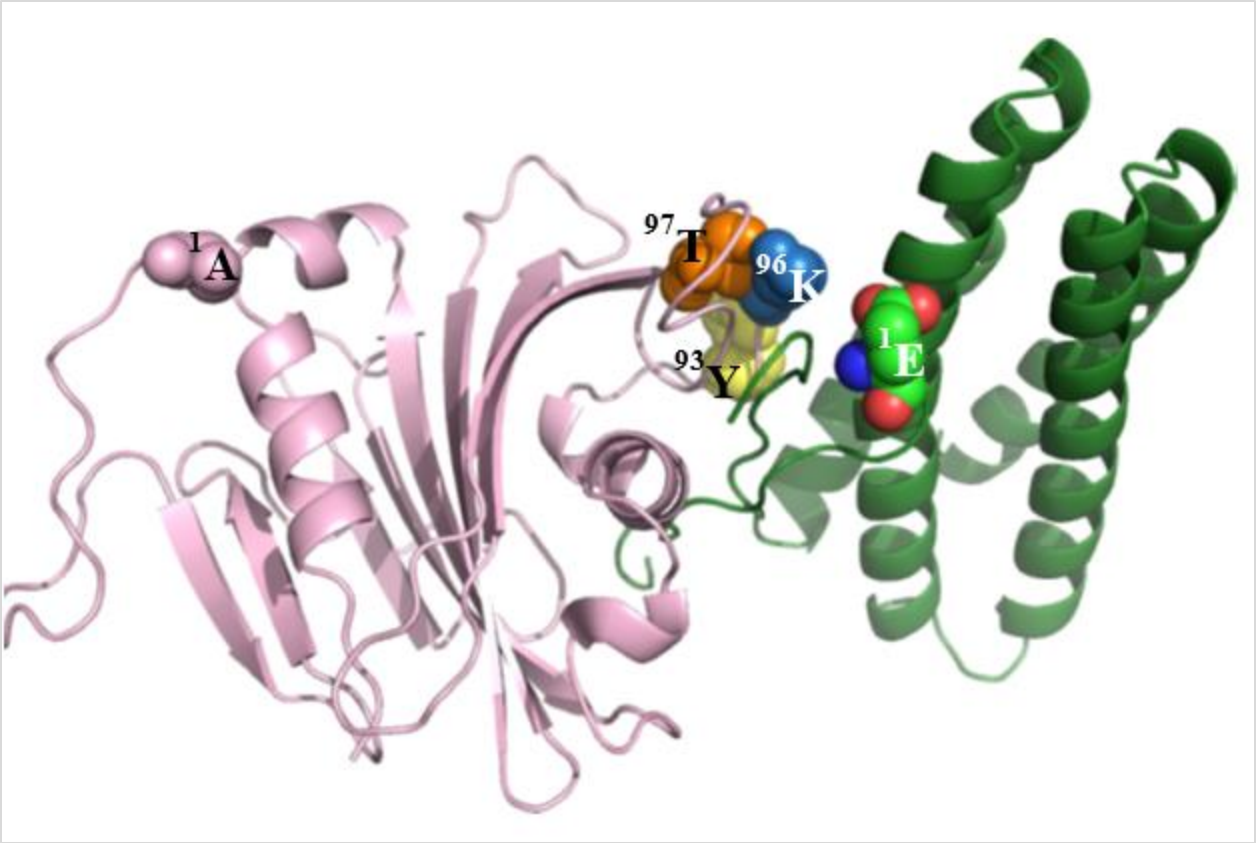


FIGURE 3.9. Interacting Domains of PsbP and PsbQ.

Schematic illustration of one possible model for BS3 crosslinked residues between PsbP and PsbQ. PsbP is shown in pink, and PsbQ in green. Residues PsbQ: ^{1}E (shown in green), PsbP: ^{93}Y, ^{96}K and ^{97}T (in yellow, blue and orange) are shown as spheres. N-terminal ^{1}A of PsbP and ^{1}E of PsbQ are labelled as pink and green spheres respectively.

The three crosslinks between PsbP and PsbQ are concurrent with both the PsbP structures with PDB IDs: 2VU4 and 4RTI. This close association of PsbQ with PsbP might explain the observation that the PsbQ helps in the stability of PsbP binding and function (Ifuku and Sato, 2002; Kakiuchi et al., 2012). This is the first report showing the interacting regions of PsbP and PsbQ in photosystem II.

3.5 Observation of a PsbQ-PsbQ Dimer

From the 44-kDa band, six crosslinked residue-pairs were identified within PsbQ (TABLE 3.1). Three of these, ^{53}K – ^{96}K, ^{128}T – ^{133}Y and ^{126}S – ^{133}Y are fully consistent with the unbound crystal structure of PsbQ (1VYK), as these residues are located 8.5 Å, 9.0 Å and 8.0 Å respectively (FIGURE 3.10 and 3.11). The other three crosslinked residue pairs are ^{125}K – ^{133}Y, ^{98}K– ^{133}Y and ^{101}K – ^{133}Y. One of these residues, ^{125}K – ^{133}Y, is 16.9 Å when mapped onto the 1VYK crystal structure shown in FIGURE 3.10. Based on the cross linker's arm-length (11.4 Å), this data suggests a possible conformational change within PsbQ occurs upon its binding to PSII.

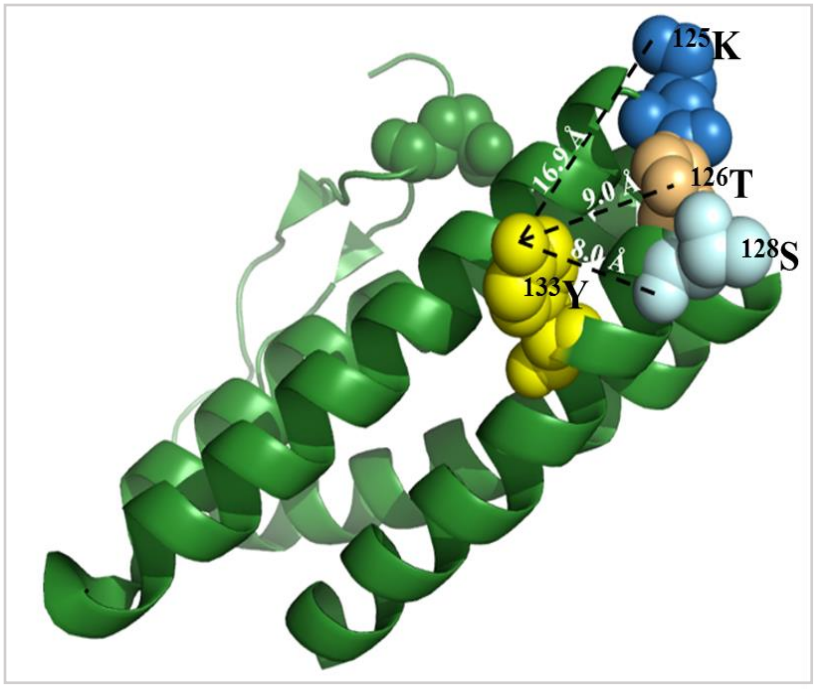


FIGURE 3.10. Intra-Molecular Crosslinked-Residues of PsbQ.
Crosslinked residue pairs are indicated. ^{133}Y is shown in yellow, ^{125}K in skyblue, ^{126}T in light-orange and ^{128}S in light-blue.

The other two crosslinked-residues ^{98}K – ^{133}Y and ^{101}K – ^{133}Y, however, are significantly more difficult to explain. In the crystal structure 1VYK, these residue pairs are located at a distance

of 33.1 Å and 30 Å, respectively (see FIGURE 3.9). The crosslinker BS3 cannot span this distance, unless an unprecedented change in the conformation of PsbQ occurs upon its binding to PSII, bringing these residues in crosslinker's active radius within a single PsbQ molecule. This is highly unlikely. Consequently, it is hypothesized, that these crosslinked residues represent interactions between two different PsbQ molecules. Previous studies (Liu et al., 2014) had identified similar results in cyanobacterial PSII, where two copies of CyanoQ (one from each monomer) were suggested to have crosslinked in their bound state to PSII. The GRAMM-X (Tovchigrechko and Vakser, 2005) and ZDOCK (Mintseris et al., 2007) programs were used to model this interaction using PsbQ (1VYK) crystal structure and the distance constraints imposed by the crosslinker ($\leq$11.4 Å). One of the possible models is shown in the FIGURE 3.10. The two PsbQ monomers associate in an antiparallel configuration, these two PsbQ copies fulfill the crosslinker length requirements, only when these two monomers are placed at an angle (not in symmetry). Thus one end of these two copies are closer to each other than the other end.

Conclusions

This chapter described important sets of results, expanding knowledge of our current understanding of PsbP and PsbQ proteins in higher plants. The method of protein crosslinking coupled to high-resolution tandem mass spectrometry is extremely useful in providing structural information in the absence of X-ray crystallography data. High mass accuracies using high resolution mass spectrometers and critical analysis of the identified ions is important to obtain near accurate or high quality structural information. The coverage of PsbP and PsbQ is greater than 90% in these experiments (see FIGURE 3.11), and the quality of spectra based on the identified ions is presented above. These proteins, PsbP and PsbQ, are important for oxygen evolution and ion-retention in PSII around the Mn_4CaO_5 cluster.

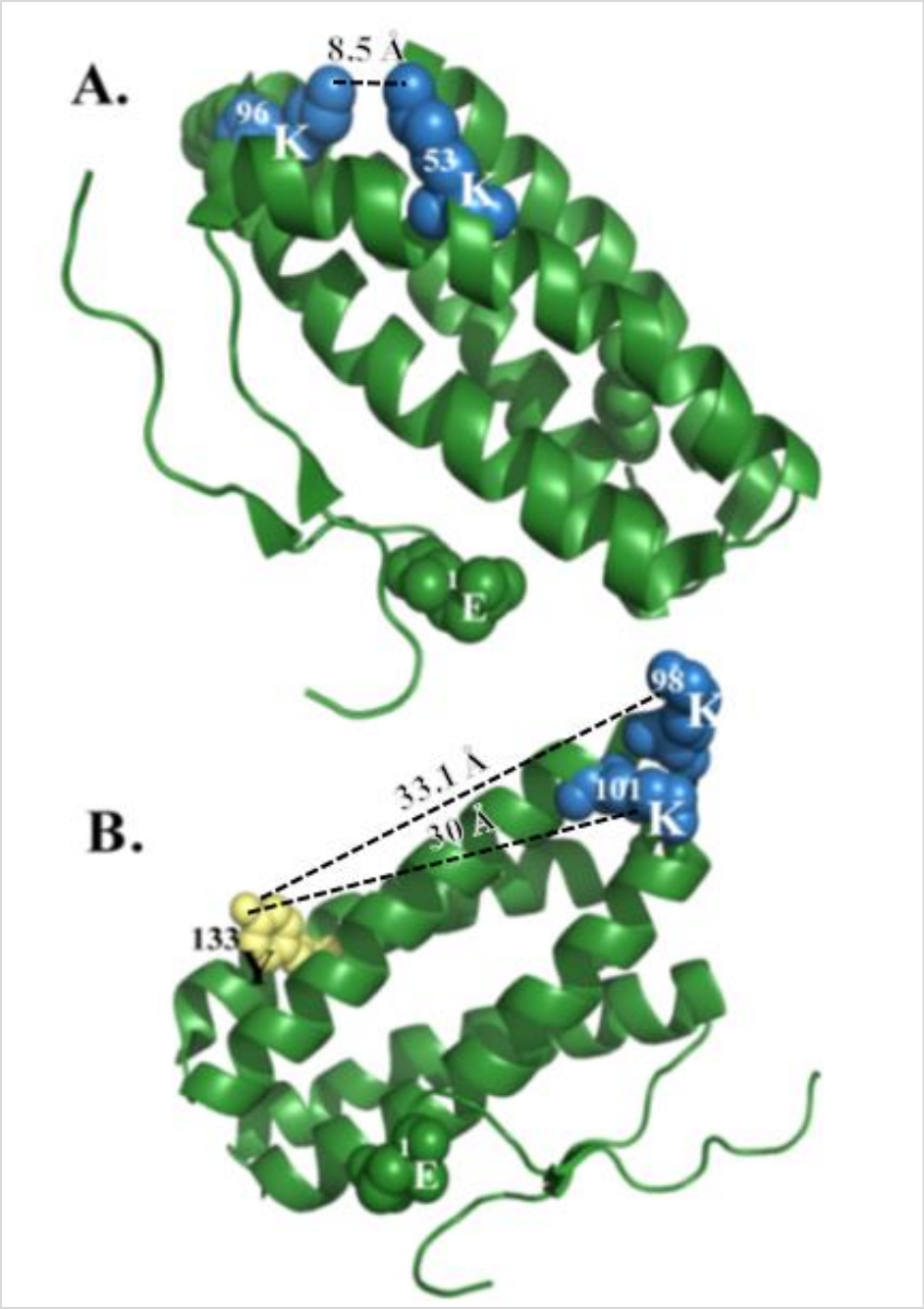
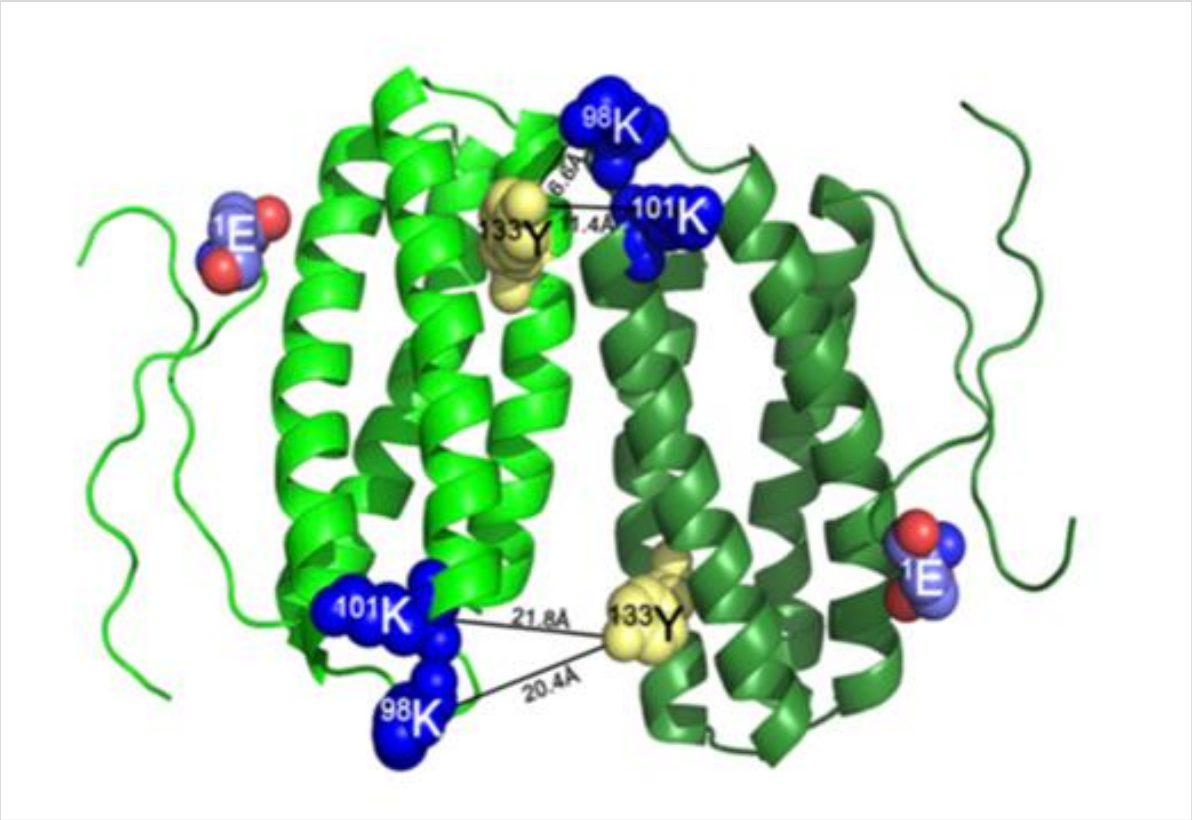


FIGURE 3.11. BS3-Crosslinked Residues of PSII Bound PsbQ.

A. Crosslinked residue-pairs ^{53}K – ^{96}K. **B.** Crosslinked residue-pairs ^{98}K – ^{133}Y and ^{101}K – ^{133}Y, mapped onto the PsbQ crystal structure, 1VYK. Lysyl- residues are shown in skyblue spheres, tyrosyl- residue in yellow and N-terminal ^{1}E in green spheres.



Mummadisetti et al., 2014

FIGURE 3.12. Interacting Domains from Two PsbQ Monomers.

A model suggesting interacting domains observed upon PSII crosslinking with BS3. The two copies of PsbQ are shown in light and dark green. The model is generated using GRAMM-X program. Lysyl residues are shown in blue and tyrosyl residues in yellow spheres. Residue PsbQ: ^{1}E is labelled for each PsbQ monomer.

Conclusions Cont…

The structure of the N-terminus of PsbP, is highly important for its binding to PSII. This structure is not determined in any of the PsbP X-ray crystal structures from different photosynthetic organisms. The results described here, identified several residues within the N-terminus of PsbP crosslinked to the C-terminus of PsbP, suggesting a non-extended N-terminus, and that it has a compact structure. Additionally these results identified two residue-pairs from PsbQ, which are > 30 Å apart on PsbQ structure, 1VYK, much farther than the crosslinker's active crosslinking radius. Hence we hypothesized that these reisudes are not from a single PsbQ, rather, these residues are contributed by two separate PsbQ copies from two PSII monomers. Protein crosslinking with BS3 also identified

a few residues in PsbP and PsbQ which are 14 Å – 16 Å apart, suggesting a conformational change within PsbP and PsbQ upon binding to PSII (based on their current unbound forms of crystal structures). Furthermore, the interacting domains of PsbP with PsbQ have also been identified which provides a framework for understanding the structural and functional interactions between these two subunits. These results will be very helpful in formulating global models of higher plant PSII.

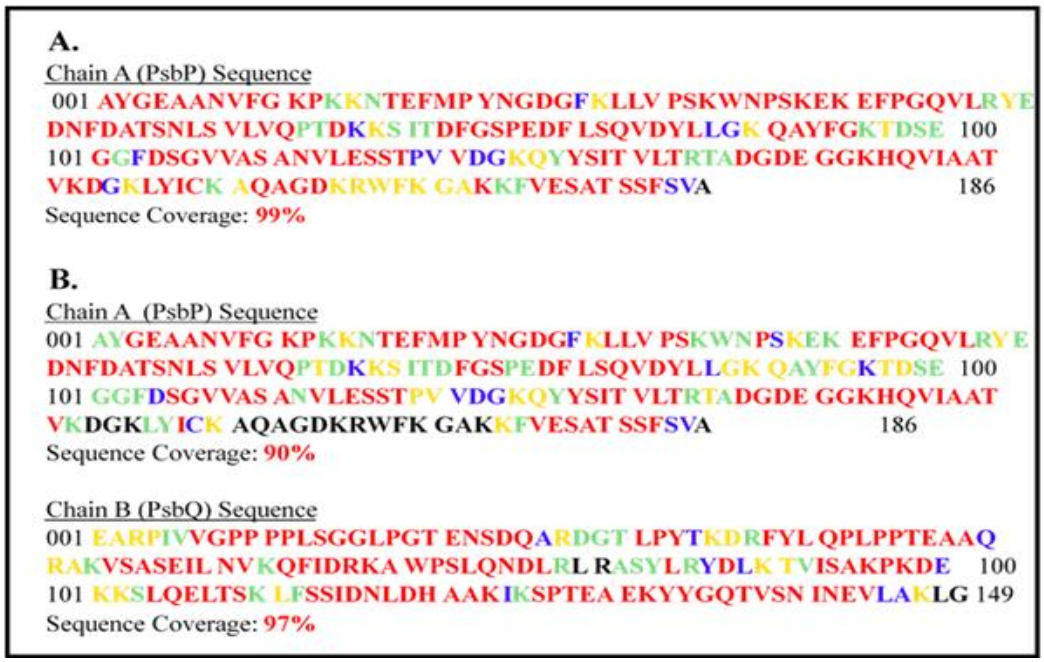


FIGURE 3.13. PsbP and PsbQ Sequences Identified by MassMatrix.

The amino acid residues coded in different colors represent sequence tags. The letters in blue, green, red or yellow indicate the identified residues in one or more peptides in a protein. Letters in blue indicate, consecutive y-ions whose mass difference equals the mass of the amino acid residue, similarly green corresponds to consecutive b ions, red corresponds to consecutive b and y ions and yellow represent the identified b or y ions which are not consecutive on a peptide.

References

Boekema, E.J., van Breemen, J.F., van Roon, H., and Dekker, J.P. (2000). Conformational changes in photosystem II supercomplexes upon removal of extrinsic subunits. Biochemistry *39*, 12907-12915.

Bricker, T.M., Roose, J.L., Fagerlund, R.D., Frankel, L.K., and Eaton-Rye, J.J. (2012). The extrinsic proteins of Photosystem II. Biochimica et Biophysica Acta *1817*, 121-142.

Cao, P., Xie, Y., Li, M., Pan, X., Zhang, H., Zhao, X., Su, X., Cheng, T., and Chang, W. (2015). Crystal structure analysis of extrinsic PsbP protein of photosystem II reveals a manganese-induced conformational change. Molecular Plant 8, 664-666.

Demetrios F. Ghanotakis, G.T.B.a.C.F.Y. (1985). On the role of water-soluble polypeptides (17,23 kDa), calcium and chloride in photosynthetic oxygen evolution. The FEBS Journal *192*, 1-3.

Eswar, N., Webb, B., Marti-Renom, M.A., Madhusudhan, M.S., Eramian, D., Shen, M.Y., Pieper, U., and Sali, A. (2007). Comparative protein structure modeling using MODELLER. Current protocols in protein science / editorial board, John E. Coligan ... [et al.] *Chapter 2*, Unit 2 9.

Ghanotakis, D.F., Topper, J.N., Babcock, G.T., and Yocum, C.F. (1984). Water-Soluble 17-Kda and 23-Kda Polypeptides Restore Oxygen Evolution Activity by Creating a High-Affinity Binding-Site for Ca-2+ on the Oxidizing Side of Photosystem-Ii. FEBS Letters *170*, 169-173.

Ghanotakis, K.K.D.F. (1991). Effect of the manganese complex on the binding of the extrinsic proteins (17, 23 and 33 kDa) of Photosystem II Photosynthesis Research *29*, 149-155.

GT, H. (1996). Bioconjugate Techniques (Academic, San Diego, CA).

Ido, K., Kakiuchi, S., Uno, C., Nishimura, T., Fukao, Y., Noguchi, T., Sato, F., and Ifuku, K. (2012). The conserved His-144 in the PsbP protein is important for the interaction between the PsbP N-terminus and the Cyt b559 subunit of photosystem II. The Journal of Biological Chemistry *287*, 26377-26387.

Ido, K., Nield, J., Fukao, Y., Nishimura, T., Sato, F., and Ifuku, K. (2014). Cross-linking evidence for multiple interactions of the PsbP and PsbQ proteins in a higher plant photosystem II supercomplex. The Journal of Biological Chemistry *289*, 20150-20157.

Ifuku, K., Nakatsu, T., Kato, H., and Sato, F. (2004). Crystal structure of the PsbP protein of photosystem II from Nicotiana tabacum. EMBO Reports *5*, 362-367.

Ifuku, K., and Sato, F. (2001). Importance of the N-terminal sequence of the extrinsic 23 kDa polypeptide in Photosystem II in ion retention in oxygen evolution. Biochimica et Biophysica Acta *1546*, 196-204.

Ifuku, K., and Sato, F. (2002). A truncated mutant of the extrinsic 23-kDa protein that absolutely requires the extrinsic 17-kDa protein for Ca2+ retention in photosystem II. Plant & Cell Physiology *43*, 1244-1249.

Kakiuchi, S., Uno, C., Ido, K., Nishimura, T., Noguchi, T., Ifuku, K., and Sato, F. (2012). The PsbQ protein stabilizes the functional binding of the PsbP protein to photosystem II in higher plants. Biochimica et Biophysica Acta *1817*, 1346-1351.

Kopecky, V., Jr., Kohoutova, J., Lapkouski, M., Hofbauerova, K., Sovova, Z., Ettrichova, O., Gonzalez-Perez, S., Dulebo, A., Kaftan, D., Smatanova, I.K., *et al.* (2012). Raman spectroscopy adds complementary detail to the high-resolution x-ray crystal structure of photosynthetic PsbP from Spinacia oleracea. PloS One *7*, e46694.

Kurowski, M.A., and Bujnicki, J.M. (2003). GeneSilico protein structure prediction meta-server. Nucleic Acids Research *31*, 3305-3307.

Kuwabara, T., Murata, T., Miyao, M., and Murata, N. (1986). Partial Degradation of the 18-Kda Protein of the Photosynthetic Oxygen-Evolving Complex - a Study of a Binding-Site. Biochimica et Biophysica Acta *850*, 146-155.

Liu, H., Zhang, H., Weisza, D.A., Vidavsky, I., Gross, M.L., and Pakrasi, H.B. (2014). MS-based cross-linking analysis reveals the location of the PsbQ protein in cyanobacterial photosystem II. Proceedings of the National Academy of Sciences of the United States of America *111*, 4638-4643.

Michoux, F., Takasaka, K., Boehm, M., Nixon, P.J., and Murray, J.W. (2010). Structure of CyanoP at 2.8 A: implications for the evolution and function of the PsbP subunit of photosystem II. Biochemistry *49*, 7411-7413.

Mintseris, J., Pierce, B., Wiehe, K., Anderson, R., Chen, R., and Weng, Z. (2007). Integrating statistical pair potentials into protein complex prediction. Proteins *69*, 511-520.

Mitsue Miyao, N.M. (1989). The mode of binding of three extrinsic proteins of 33 kDa, 23 kDa and 18 kDa in the photosystem II complex of spinach. biochim Biophys Acta-Bioenergetics *3*, 315-321.

Miyao, M., Fujimura, Y., and Murata, N. (1988). Partial Degradation of the Extrinsic 23-Kda Protein of the Photosystem-Ii Complex of Spinach. Biochimica et Biophysica Acta *936*, 465-474.

Miyao, M., and Murata, N. (1984). Calcium ions can be substituted for the 24-kDa polypeptide in photosynthetic oxygen evolution. FEBS Letters *168*, 118-120.

Mummadisetti, M.P., Frankel, L.K., Bellamy, H.D., Sallans, L., Goettert, J.S., Brylinski, M., Limbach, P.A., and Bricker, T.M. (2014). Use of protein cross-linking and radiolytic footprinting to elucidate PsbP and PsbQ interactions within higher plant Photosystem II. Proceedings of the National Academy of Sciences of the United States of America *111*, 16178-16183.

Nishimura, T., Nagao, R., Noguchi, T., Nield, J., Sato, F., and Ifuku, K. (2016). The N-terminal sequence of the extrinsic PsbP protein modulates the redox potential of Cyt b559 in photosystem II. Scientific Reports *6*, 21490.

Roose, J.L., Frankel, L.K., and Bricker, T.M. (2010). Documentation of significant electron transport defects on the reducing side of photosystem II upon removal of the PsbP and PsbQ extrinsic proteins. Biochemistry *49*, 36-41.

Sinz, A. (2006). Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions. Mass Spectrometry Reviews *25*, 663-682.

Tomita, M., Ifuku, K., Sato, F., and Noguchi, T. (2009). FTIR evidence that the PsbP extrinsic protein induces protein conformational changes around the oxygen-evolving Mn cluster in photosystem II. Biochemistry *48*, 6318-6325.

Tovchigrechko, A., and Vakser, I.A. (2005). Development and testing of an automated approach to protein docking. Proteins *60*, 296-301.

Yi, X., Hargett, S.R., Liu, H., Frankel, L.K., and Bricker, T.M. (2007). The PsbP protein is required for photosystem II complex assembly/stability and photoautotrophy in Arabidopsis thaliana. The Journal of Biological Chemistry *282*, 24833-24841.

CHAPTER 4.

USE OF PROTEIN CROSSLINKING AND TANDEM MASS SPECTROMETRY TO ELUCIDATE THE STRUCTURE OF PsbO WITHIN HIGHER PLANT PSII

4.1 Introduction

The water oxidation machinery of PSII (OEC) is an ensemble of a Mn_4CaO_5 cluster along with the ligands provided by intrinsic membrane proteins. PsbO is essential for the stability of this cluster and ion retention process at the OEC (Bricker and Frankel, 1998). This protein is uniformly present in all oxygenic organisms along with PsbU, PsbV, CyanoP and CyanoQ (in the cyanobacteria), or with PsbP, PsbQ and PsbR (in higher plants). In red algae, extrinsic proteins consist of PsbO, PsbU, PsbV and PsbQ′, diatoms, however, have all the red algal extrinsic proteins along with Psb31 (Okumura et al., 2008).

Until recently, the crystal structures of PSII have been available only from thermophilic cyanobacteria (Ferreira et al., 2004; Guskov et al., 2009; Kamiya and Shen, 2003; Loll et al., 2005; Suga et al., 2015; Umena et al., 2011; Zouni et al., 2001). Recently, a red algal PSII crystal structure from *Cyanidium caldarium* has been presented with a resolution of 2.76 Å (Ago et al., 2016). Although red algae are eukaryotes, they do not lie on the green plant lineage. Our current understanding of the organization of the extrinsic proteins from higher plant has principally been from the cyanobacterial crystal structure, and thus interactions of PsbP, PsbQ and PsbR proteins, which are not present in cyanobacteria, with PSII are not well understood. The structure of higher plant PsbO has been approximated based on the structure of cyanobacterial PsbO, which is present in the PSII crystal structure. In higher plants, biochemical evidence suggests a second copy of PsbO per PSII monomer (more details in Chapter 1, Introduction). With the lack of higher plant PsbO

crystal structure and a sequence identity of only ~40% between cyanobacterial and higher plant (spinach) forms of PsbO, studies relating directly to higher plant PsbO are required.

In this section, we have used protein crosslinking by EDC coupled to tandem mass spectrometry to elucidate the structure of higher plant PsbO in its bound form to PSII. Twenty-four intramolecular crosslinked products were identified within the PsbO protein. These experiments allowed the modeling of the N-terminus of higher plant PsbO, which contains a ten amino-acid extension not seen in cyanobacterial PsbO. Subsequently, it helped in identifying the structure and location of the two binding determinants required for its complete binding to PSII region (Popelkova et al., 2002a; Popelkova et al., 2002b).

4.2 Results

In our experiment, PSII membranes from fresh market spinach were obtained and suspended at 2mg Chl/mL in the SMN buffer and frozen at -80˚C until use. Protein crosslinking was performed using a crosslinker concentration of 6.25 mM of EDC with 5 mM of N-hydroxy sulfo succinamide (sulfo-NHS) in the crosslinking buffer. The reaction was carried out in the dark, at room temperature for two hours with occasional shaking of the vial. The reaction was quenched using 100 mM ammonium bicarbonate and incubation for 20 minutes at room temperature (Refer Chapter 2 on Materials and Methods for details). The crosslinked PSII membranes were initially washed with 1.0 M NaCl to remove PsbP, PsbQ and any crosslinked forms of these subunits not crosslinked to either the intrinsic membrane proteins or PsbO. Immunoblotting of the soluble fraction from the 1.0 M NaCl extracts, did not yield any crosslinked products (data not shown). The NaCl-washed PSII membranes were then washed with 1.0 M $CaCl_2$ in SMN buffer. The soluble fraction from these membranes were analyzed by LiDS-PAGE following by immunoblot analysis.

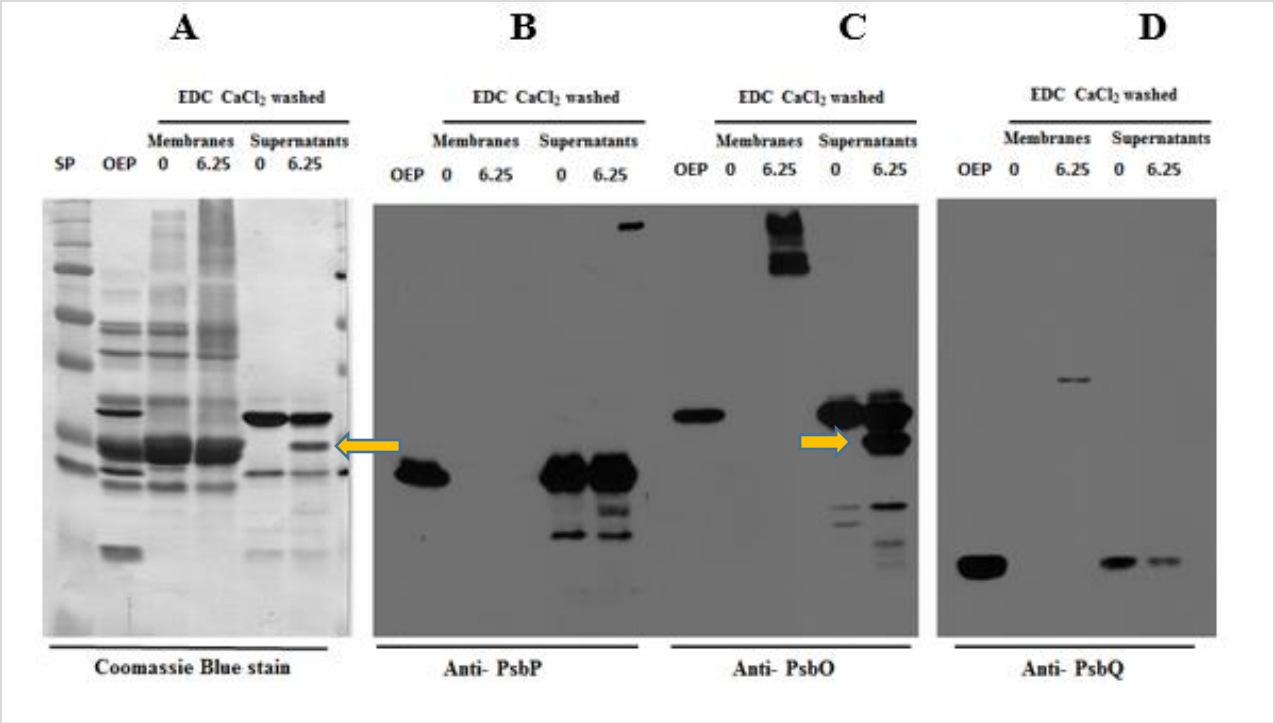


FIGURE 4.1. EDC Crosslinking of PsbO bound to PSII.

The crosslinking was performed using the protocols described in the Chapter 2, Materials and Methods. Illusterated are the immunoblots from 1.0 M $CaCl_2$ extracts obtained by washing crosslinked PSII membranes (after intial washing with 1.0 M NaCl, immunoblots from NaCl did not yield any crosslinked product, and hence are not shown here). From left to right: SP refers to standard proteins, OEP or Oxygen evolving preparations are the intact PSII, crosslinked membranes or supernatants at 0 mM and 6.25 mM concentration of EDC. The first panel was stained with Coomassie Blue, and the next three lanes with antibodies against PsbP, PsbO and PsbQ, labelled at the bottom of the figure.

The FIGURE 4.1 illustrates the analysis of these CaCl2-washed membranes and supernatants by LiDS-PAGE and "Western" blotting. Panel A, was stained with Coomassie blue after transferring the proteins onto a PVDF membrane. The rest of the panels were probed with either α-PsbP, α-PsbO or α-PsbQ antibodies. EDC crosslinker concentrations are represented as 0 mM (control) and 6.25 mM (for EDC + Sulfo-NHS). SP refers to the standard proteins.

Protein crosslinking identified several crosslinked products. The α-PsbO and α-PsbQ antibodies detected bands in membrane lanes in addition to the supernatant lanes. The extrinsic proteins (PsbO, PsbP and PsbQ) can be extracted from PSII under high salt conditions, 1M $CaCl_2$ in this case. The "Western" blots also identified a protein band (orange arrows) that migrates between PsbP and PsbO proteins, not seen in the control lane. This band had an apparent molecular mass of 25 kDa and is only seen in the panel probed with α-PsbO (and not in α-PsbP or α-PsbQ). Since this band migrates below PsbO and is detected only by α-PsbO, it is either a degraded PsbO or a crosslinked product of PsbO. The protein yield of this band was high (μg levels), and is also seen as a thick band on coomassie-stained panel, hence it was chosen for further analysis by tandem mass spectrometry.

For Mass Spectrometry, the 25 kDa band was excised from the Coomassie Blue stained gel, and subjected to proteolytic digestion with either Trypsin or Trypsin + Lys C. These enzymes digest the proteins at the C-terminal end of either arginine or lysine, unless either is followed by a proline. Post digestion, the peptides were extracted from the protein bands, which were then sent through a C18 ZipTip® to concentrate the sample, purify it from any gel fragments or left over Coomassie Blue stain and for better separation of the peptides in HPLC. The concentrated peptide samples were sent through a reverse phase column followed by tandem mass spectrometry or MS/MS (ESI- LTQ FT). Mass spectrometric analysis indicated that the 25 kDa band was indeed the PsbO protein, with 100% coverage, and complete coverage of C- and N-terminus of the PsbO sequence, suggesting no degradation of the protein. Thus, 25 kDa appears to result from internal crosslinking of PsbO with EDC.

4.3 Identification and Analysis of Crosslinked Products

The mass modifications were analyzed by two programs, MassMatrix online search engine (Xu and Freitas, 2009), and StavroX, ver. 3.4 (Gotze et al., 2012). A FASTA library, containing the proteins PsbO, PsbP and PsbQ, and the decoy library containing the same proteins, however, with their sequences reversed (to avoid false positives), were used for crosslinks search. Criteria of *P* values $\leq$0.0001 were used for MassMatrix searches of crosslinked peptides. In case of StavroX however, a non-probabilistic scoring system, the peptides were searched for a score >100. Additionally the scores of the crosslinked residues greater than or equal to 3*(FDR) were further evaluated. FDR corresponds to a score at which false discovery is less than 5%, which means that the chances of occurrence is $\geq$95.5%. And, a score of hundred is suggested to have a false discovery of less than 2% (Gotze et al., 2012).

The crosslinked residues from the 25 kDa band identified twenty-four intra-chain crosslinked products within PsbO. The crosslinks identified from MassMatrix program demonstrated very low *P* values ranging from 1×10^{-4} to 5×10^{-10}. The TABLE 4.1 summarizes the crosslinked products identified using MassMatrix and StavroX. For most of the crosslinked products observed, analogous residue pairs are present in cyanobacterial PsbO, and are shown in the last column of the TABLE 4.1. As expected, these residues from cyanobacterial PsbO had their Cα-Cα distances within 12 Å, further suggesting the structural similarity between cyanobacterial and spinach PsbO supported by our experimental data.

TABLE 4.1. EDC-Crosslinked Residues from PSII-Bound PsbO.

Crosslinked residue-pairs identified in $CaCl_2$ extracted PsbO from crosslinked spinach PSII membranes, by MassMatrix and StavroX programs. Analogous residues from cyanobacteria are indicated.

PsbO Crosslinks	StavroX: [1]Score	[2]FDR	MassMatrix: [3]P value	[4]Crystal Structure Distance (C_α-C_α)
^{4}K -^{36}D	--	--	6.3×10^{-5}	[5]NA
^{9}D -^{14}K	172	52	5.0×10^{-7}	NA
^{10}E -^{14}K	182	52	3.1×10^{-5}	NA
^{14}K -^{18}E	253	59	2.5×10^{-7}	8.1 Å
^{60}K -^{62}E	--	--	1.0×10^{-8}	5.8 Å
^{60}K -^{234}D	278	59	2.5×10^{-9}	6.9 Å
^{62}E -^{66}K	--	--	1.2×10^{-8}	11.5 Å
^{62}E -^{233}K	--	--	5.0×10^{-7}	9.9 Å
^{66}K -^{71}D	--	--	6.3×10^{-7}	10.8 Å
^{66}K -^{234}D	--	--	5.0×10^{-9}	16.1 Å
^{86}D -^{105}K	204	57	--	12.2 Å
^{89}E -^{105}K	221	59	--	7.2 Å
^{97}D -^{101}K	183	59	6.3×10^{-8}	9.5 Å
^{101}K -^{104}E	--	--	1.0×10^{-10}	9.7 Å
^{103}E -^{105}K	--	--	1.9×10^{-9}	6.9 Å
^{105}K -^{106}D	--	--	3.1×10^{-5}	3.8 Å
^{137}K -^{139}E	207	59	--	9.9 Å
^{137}K -^{144}D	207	59	--	9.0 Å
^{180}D -^{186}K	224	59	--	NA
^{180}D -^{190}K	--	--	1.0×10^{-4}	NA
^{181}E -^{186}K	223	59	--	9.7 Å
^{181}E -^{190}K	180	57	--	14.4 Å
^{182}E -^{186}K	202	59	--	6.2 Å
^{183}E -^{186}K	202	59	--	5.3 Å

[1]A significant score: qualifying crosslinks have a StavroX score >100 and >3x FDR.
[2] FDR, False Discovery Rate: value represents a score at which false discovery is less than 5%
[3] P value; calculated by MassMatrix using max (pp1, pp2), pptag.
[4] Distances between the corresponding residues in the cyanobacterial PsbO structure of PS II.
[5] Not applicable, as no corresponding residues are present.

The FIGURE 4.2 (A) illustrates cyanobacterial PsbO, PDB: 3WU2 (orange) aligned to spinach PsbO, threaded using cyanobacterial structure as a template. The domains A through E represent the regions where crosslinked residues were identified, and apparently these domains (specially A, B, D and E) as seen in FIGURE 4.2 (A) exhibit most differences in the cyanobacterial (orange) versus higher plant (green) PsbO. The figure 4.2 (B) illustrates the crosslinked residues as spheres shown in different colors for each domain (A through E). The domain A corresponds to the residues at the N-terminus of higher plant PsbO (^{1}E-^{10}E), not present in the cyanobacterial PsbO. Three crosslinked residues were identified in this domain. Additionally, higher plant PsbO lacks an insertion of eight residues seen in cyanobacteria, known as cyano-loop (De Las Rivas and Barber, 2004), crosslinks were idenfied in the analogous region in spinach near ^{138}P-^{139}E random coil region, labeled domain 'B'. Also, β-barrel region, near domain 'B', idenfied several crosslinked residues labeled as 'C'. Furthermore, crosslinked residues were identified in the region ^{178}R-^{180}D, labeled 'E', a domain missing in cyanobacterial PsbO. This information placed strong distance-constraints on the probable structures of PsbO and was helpful in proposing higher plant PsbO models.

The FIGURE 4.3 illustrates a representative spectrum, along with the heat map and its b- and y- type ions for the peptide ^{6}L – ^{20}K of PsbO indicating the quality of the data presented in the TABLE 4.1. This peptide had an intra-molecular crosslink between the residues ^{14}K and ^{18}E. The crosslinked residue-pair has a median *P* value of 2.5 x 10^{-7}. As shown in the figure, we obtained a complete b and y ion series. The gap in the ion series is due to intra-peptide crosslink formation. The intra-peptide crosslink generates a loop in the peptide that is not detected by any of the currently available softwares (Xu and Freitas, 2009).

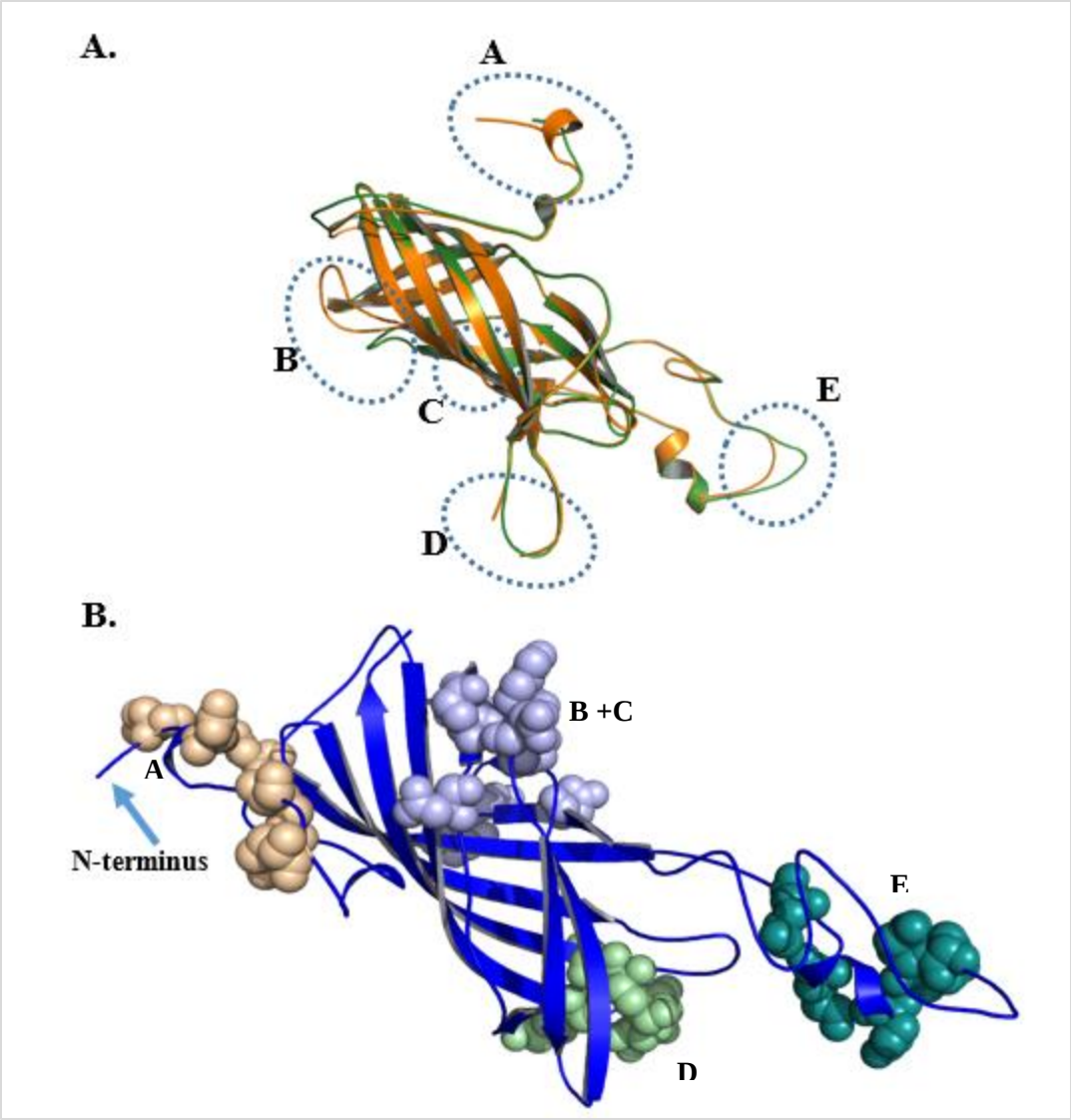
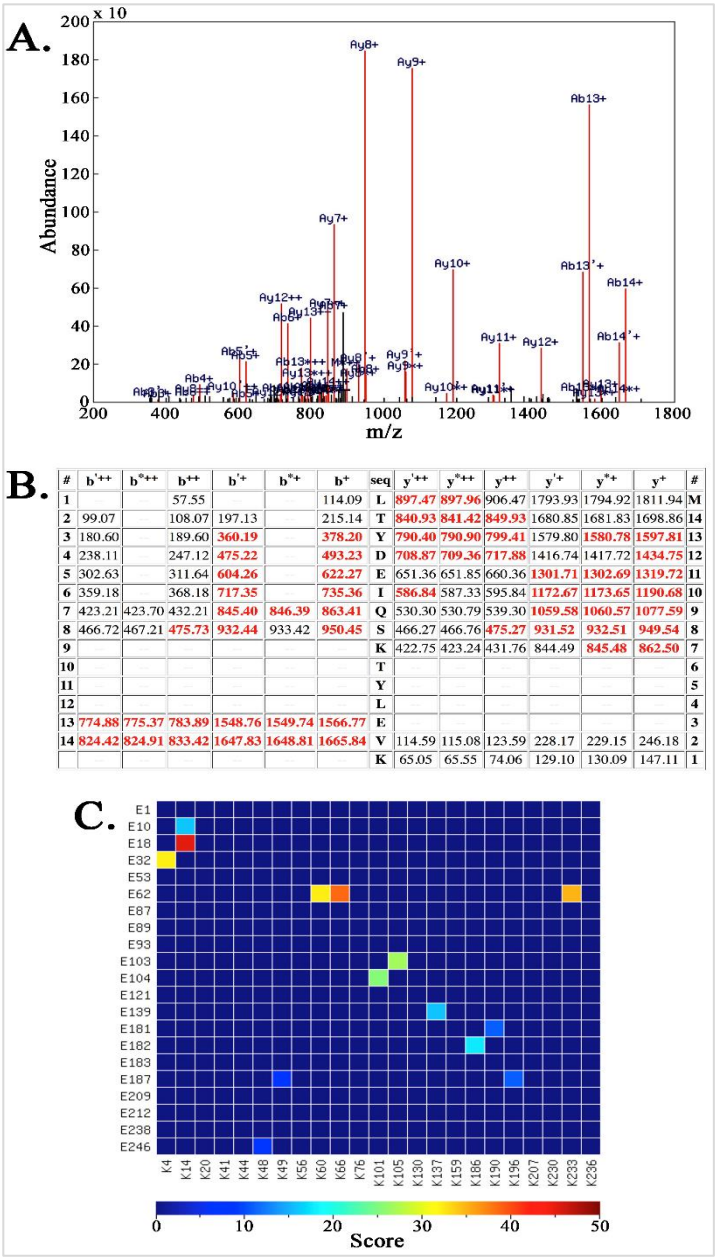


FIGURE 4.2. PsbO from Spinach and Cyanobacteria and Crosslinked Residues Identified in four Domains.

A. X-ray crystal structure of PsbO from *T.vulcanus* is aligned to the spinach PsbO. B. The crosslinked residues were identified in four separate domains. These crosslinked-residues are shown as spheres on one of the top modelled PsbO incorporating crosslinking-residues distance constraints. The crosslinked residues at the N-terminal region, extended loop region and cyano loop are the most interesting. Different colored spheres indicate that the crosslinked-residues were clustered in those domains.



#	b'++	b*++	b++	b'+	b*+	b+	seq	y'++	y*++	y++	y'+	y*+	y+	#
1			57.55			114.09	L	897.47	897.96	906.47	1793.93	1794.92	1811.94	M
2	99.07		108.07	197.13		215.14	T	840.93	841.42	849.93	1680.85	1681.83	1698.86	14
3	180.60		189.60	360.19		378.20	Y	790.40	790.90	799.41	1579.80	1580.78	1597.81	13
4	238.11		247.12	475.22		493.23	D	708.87	709.36	717.88	1416.74	1417.72	1434.75	12
5	302.63		311.64	604.26		622.27	E	651.36	651.85	660.36	1301.71	1302.69	1319.72	11
6	359.18		368.18	717.35		735.36	I	586.84	587.33	595.84	1172.67	1173.65	1190.68	10
7	423.21	423.70	432.21	845.40	846.39	863.41	Q	530.30	530.79	539.30	1059.58	1060.57	1077.59	9
8	466.72	467.21	475.73	932.44	933.42	950.45	S	466.27	466.76	475.27	931.52	932.51	949.54	8
9							K	422.75	423.24	431.76	844.49	845.48	862.50	7
10							T							6
11							Y							5
12							L							4
13	774.88	775.37	783.89	1548.76	1549.74	1566.77	E							3
14	824.42	824.91	833.42	1647.83	1648.81	1665.84	V	114.59	115.08	123.59	228.17	229.15	246.18	2
							K	65.05	65.55	74.06	129.10	130.09	147.11	1

FIGURE 4.3. Mass Spectrometry Analysis for the Crosslinked-Residues ^{14}K – ^{18}K.

Shown is the mass spectrometry data obtained for the median P value crosslinked product ^{14}K – ^{18}E (TABLE 4.1, P value = 2.5 x 10^{-7}).
A. Mass spectra obtained for this peptide. Identified ions are shown in red, those not identified in black. **B**. Table of predicted ions from this peptide. Identified ions are shown in red, those not identified in black. MassMatrix does not identify residues lying between the crosslinked residues. Note the nearly complete y- and b-ion series obtained. **C.** Heat map which assists in the identification of crosslinked species. The identified peptide is shown in red.

4.4 Protein Modeling

The Clustal Omega (Sievers et al., 2011) sequence alignments for PsbO from spinach and *T. vulcanus* identified a sequence identity of 47% and a sequence similarity of ~65 % between the two organisms. Based on their alignment, spinach PsbO structure was threaded by a homology-modelling server called SWISS-MODEL (Arnold et al., 2006), using cyanobacterial PsbO as a template. The secondary structure analysis was performed using the program Genesilico Metaserver (Kurowski and Bujnicki, 2003) which utilizes eighteen independent algorithms and provides secondary-structure with each algorithms in addition to their combined consensus. PsbO models were generated using the program MODELLER (Eswar et al., 2007), customized to model the N-terminus based on the consensus secondary structure prediction obtained from the Genesilico Metaserver and the distance constraints for all twenty-four crosslinked-residues obtained by EDC crosslinking. Please note that the higher plant PsbO has a ten amino-acid extension at its N-terminus, which is not found in cyanobacterial PsbO. Hence the consensus secondary structure is used for its modeling which suggested a short helix between ^{7}T-^{10}E at the N-terminus of PsbO.

EDC crosslinker forms an amide bond between the primary amines and the carboxylates within the van der Waals distance (Grabarek and Gergely, 1990; Kunkel et al., 1981). These interactions were modelled by constraining the Cα-Cα distances upto 12 Å (Rozbesky et al., 2012; Sriswasdi et al., 2014) shown in the FIGURE 4.4. A total of twenty energy-minimized models were generated using the program MODELLER. These structures were optimized or energy-minimized by molecular dynamic refinement protocols either in the absence or presence of the distance constraints (with or without the N-terminal predicted short-helix). From these, the top ten, lowest free-energy models were selected. The FIGURE 4.4 (A) illustrates ten lowest-energy models without any distance constraints obtained from the crosslinking data, FIGURE 4.4 (B) illustrates the models

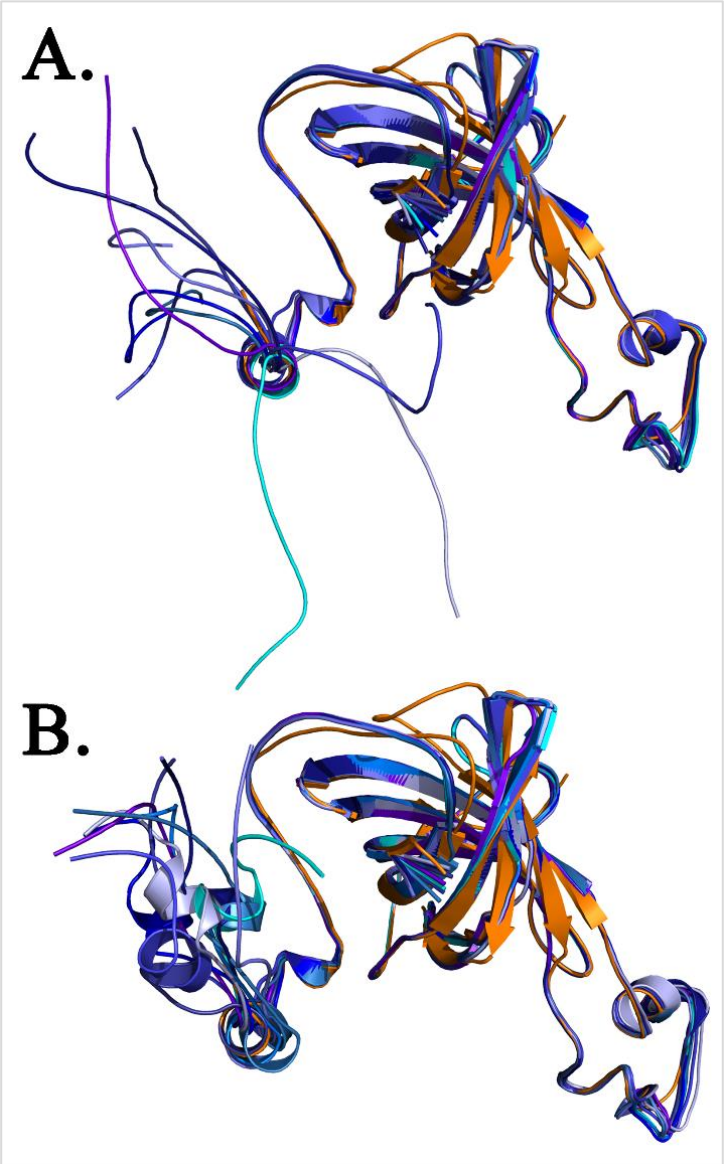


FIGURE 4.4. Spinach PsbO Modelled using MODELLER.
Shown are the top 10 models for A. No EDC-crosslinking distance constraints incorporated, and B. with the incorporation of EDC-distance constraints. The individual PsbO models are shown in various shades of blue and these are aligned with cyanobacterial PsbO, shown in orange. All models exhibited similar low DOPE scores of about -19,500.

incorporating the distance constraints and N-terminal short helix. The stereochemistry of these models were assessed by Ramachandran analysis using the programs PROCHECK (Sriswasdi et al., 2014) and RAMPAGE (Lovell et al., 2003). The FIGURES 4.5 - 4.7 illustrate the Ramachandran

analysis for the threaded model of spinach PsbO obtained from cyanobacterial template, cyanobacterial PsbO (PDB: 3WU2) and distance-constrained spinach PsbO model with the median DOPE score respectively. All of our models qualified the Ramachandran analyses. The Ramachandran analyses requires the protein models to have 90% of their residues in the favored regions or less than 2% of the residues in the unfavored regions. The PsbO models with the crosslinking distance constraints showed similar Ramachandran [φ, ψ] plots in comparison to the cyanobacterial PsbO. The threaded structure of spinach PsbO contained three residues in the disallowed region, however the spinach PsbO incorporating EDC crosslinking data contained only one residue. From structural validation perspective, this suggests that our crosslinking data added constraints that allowed PsbO to attain a conformation with the most stable and feasible Φ and φ angles. More details on Ramachandran analysis is described in CHAPTER 1.

4.5 Discussion

In this section, new information concerning the structure of higher plant PsbO is provided. As mentioned earlier, the higher plant crystal structure is currently unavailable; hence cyanobacterial PsbO is a good starting point to study PsbO from higher plants. The sequence similarity suggests a similar overall fold of PsbO in cyanobacteria and higher plants. The sequence alignments indicated an extension or a deletion of certain domains. Protein threading and homology modeling tools helped in obtaining a model for spinach PsbO from its cyanobacterial counterpart (as a template) and this spinach PsbO model was used as a reference for crosslinking studies. Protein crosslinking on intact PSII membranes identified a 19 kDa band on LiDS PAGE. Mass spectrometric analysis of this protein band identified twenty-four crosslinked-residues scattered in four domains of PsbO, clustered in both the beta-barrel and the loop regions including the ten-residue extension at the N-terminus of spinach PsbO (absent in cyanobacterial form), and in a region analogous to cyano-loop.

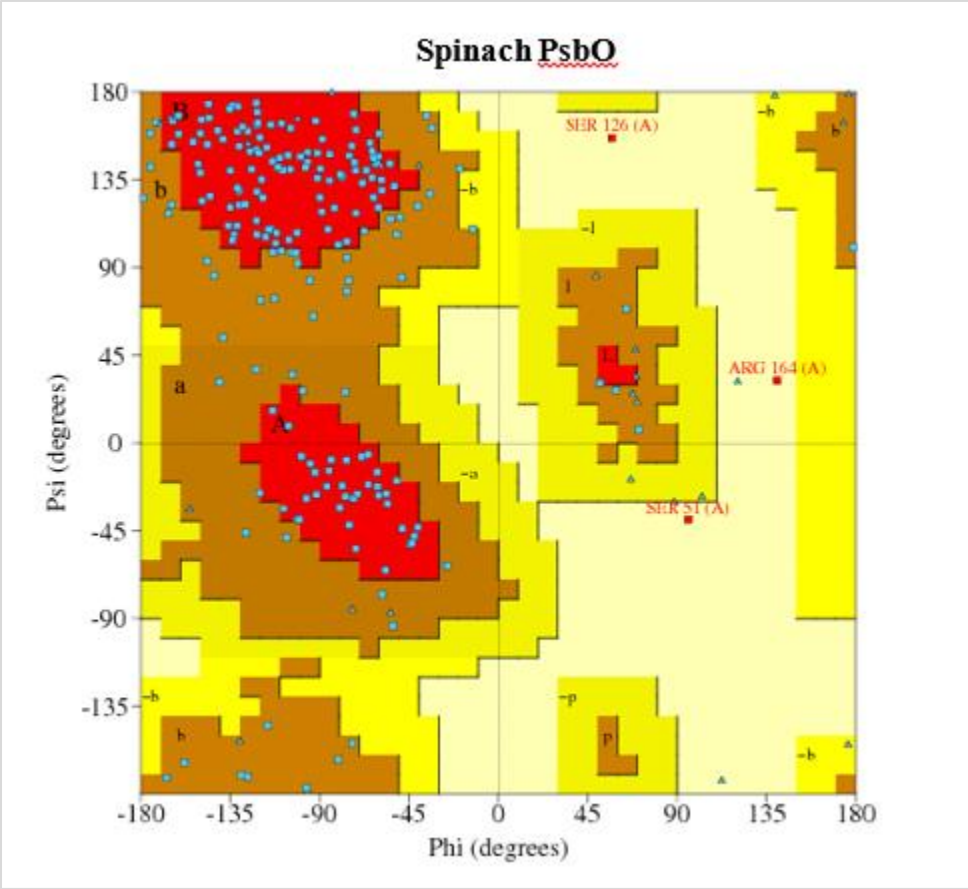


FIGURE 4.5. Ramachandran Analysis of Spinach PsbO Threaded from Cyanobacterial PsbO.

Energetically allowed regions of backbone diahedral angles are shown in the figure. All non-glycine amino acid residues are shown as small blue squares, and glycines as triangles. Residues within yellow, brown or red regions are considered allowed, and those in the background (light yellow region) are disallowed. Residues in this regions are labelled. Epsilon (P) is used for the ϕ/ψ combinations associated with a short turns (about 1-2 residues) between α helices or β strands (Morris et al., 1992).

A - Core alpha
a - Allowed alpha
~a - Generous alpha
B - Core beta
b - Allowed beta
~b - Generous beta
L - Core left-handed alpha
l - Allowed left-handed alpha
~l - Generous left-handed alpha
p - Allowed epsilon
~p - Generous epsilon

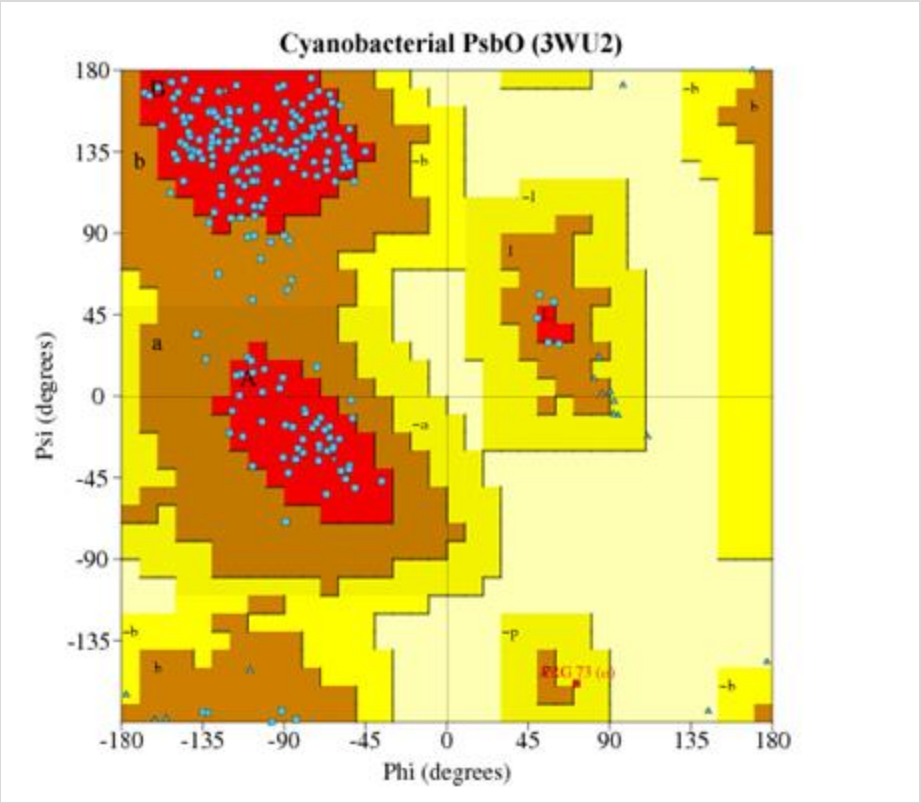


FIGURE 4.6. Ramachandran Analysis of Cyanobacterial PsbO 3WU2.

Energetically allowed regions of backbone diahedral angles are shown in the figure. All non-glycine amino acid residues are shown as small blue squares, and glycines as triangles. Residues within yellow, brown or red regions are considered allowed, and those in the background (light yellow region) are disallowed. Residues in this regions are labelled. Epsilon (P) is used for the ϕ/ψ combinations associated with a short turns (about 1-2 residues) between α helices or β strands (Morris et al., 1992).

A - Core alpha	**L** - Core left-handed alpha
a - Allowed alpha	**l** - Allowed left-handed alpha
~a - Generous alpha	**~l** - Generous left-handed alpha
B - Core beta	**p** - Allowed epsilon
b - Allowed beta	**~p** - Generous epsilon
~b - Generous beta	

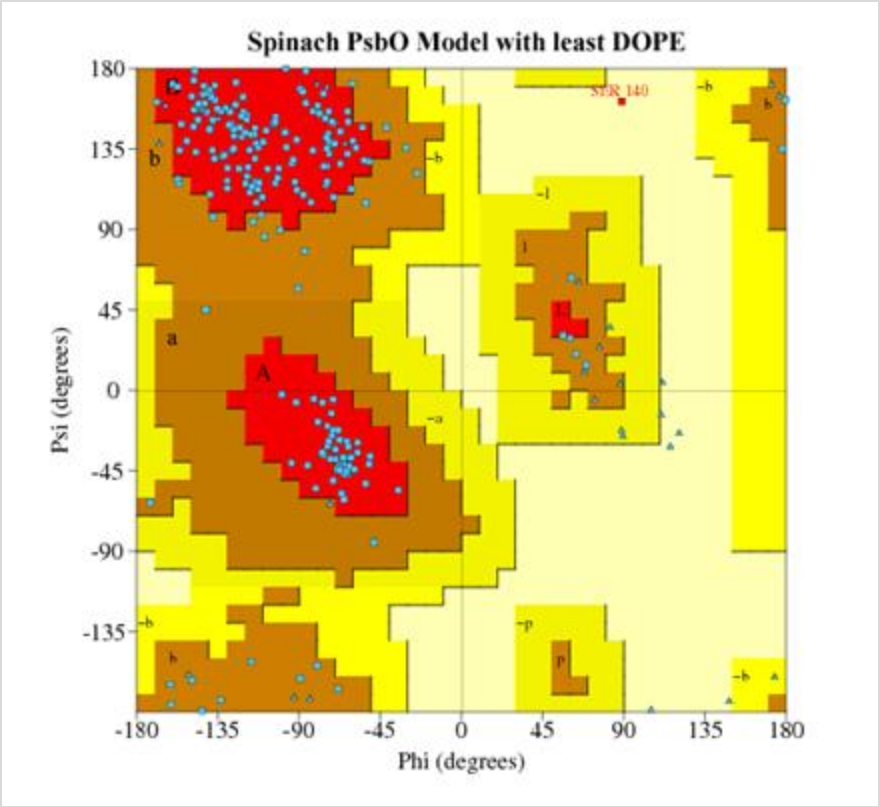


FIGURE 4.7. Ramachandran Analysis of Spinach Model with the Median DOPE Score.

Energetically allowed regions of backbone diahedral angles are shown in the figure. All non-glycine amino acid residues are shown as small blue squares, and glycines as triangles. Residues within yellow, brown or red regions are considered allowed, and those in the background (light yellow region) are disallowed. Residues in this regions are labelled. Epsilon (P) is used for the ϕ/ψ combinations associated with a short turns (about 1-2 residues) between α helices or β strands (Morris et al., 1992).

A - Core alpha
a - Allowed alpha
~a - Generous alpha
B - Core beta
b - Allowed beta
~b - Generous beta
L - Core left-handed alpha
l - Allowed left-handed alpha
~l - Generous left-handed alpha
p - Allowed epsilon
~p - Generous epsilon

Discussion Contd…The N-terminus of higher plant PsbO is particularly interesting since this domain is missing in cyano PsbO. We identified three crosslinked-residues in this domain. The PsbO N-terminus illustrated in FIGURE 4.8 identified two short helices. Thus PsbO doesn't have an extended N-terminus; rather, it is a much compact structure.

EDC crosslinking also identified other bands that were only seen in EDC crosslinked lanes and not found in the control lanes of the "Western" blot. These bands were (1) in α-PsbP supernatant lane, a protein band migrated below PsbP. (2) in α-PsbO membranes, a band (more like a smear) migrated much higher than PsbO, suggesting that PsbO is crosslinked to membrane proteins in a supercomplex. (3) in α-PsbQ membrane lane, suggesting that PsbQ was crosslinked to the membrane proteins. The complexity and concentration of these products could not yield any result.

Sequence alignments from *T.vulcanus* and *Spinacea olearacea* are shown in FIGURE 4.8 (A). The two binding determinants described earlier (Popelkova and Yocum et al, 2002) are shown in pink and green boxes, and are shown as dotted circles in FIGURE 4.8 (B). Domain I is present in both cyanobacteria and higher plants however Domain II is specific to higher plants. Deletion of domain I or II led to only 50% binding of PsbO to PSII, and only 40% oxygen evolution compared to wild type. And hence complete binding of PsbO to PSII required the presence of both the domains.

Biochemical evidence suggests presence of two copies of PsbO per PSII (refer PsbO in Chapter I). It is hypothesized that these domains could be region for the binding of second copy of PsbO in higher plants. With EDC crosslinking, we identified the structure of the N-terminus and located these binding determinants to the short helices at the N-terminus. In the absence of a crystal structure for higher plant PsbO, our PsbO serves as a low-resolution structure. Further studies are necessary to identify the location of second copy of PsbO (based on biochemical/biophysical analysis) and the interacting domains of this protein with other extrinsic proteins. Biochemical analysis suggests a sequential binding of PsbP and PsbQ only after PsbO binds to PSII in spinach. Other studies have shown that PsbP or PsbQ interacts with PsbO. These experiments have to be taken further to identify the locations of all the extrinsic proteins that entirely complement the PSII structure.

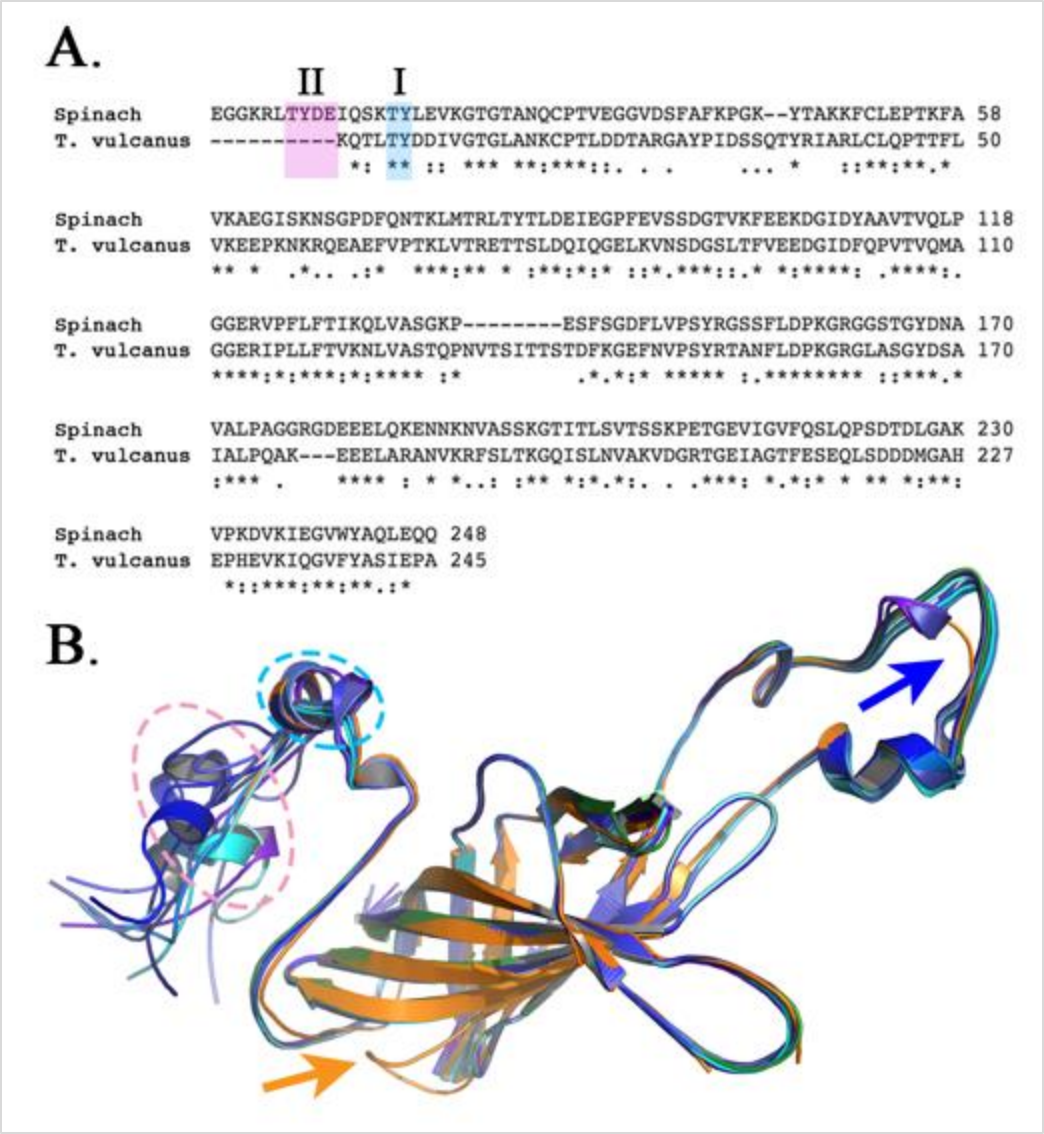


FIGURE 4.8. Binding Determinants of Higher Plant PsbO.

A. Clustal Omega alignment of spinach and *Thermosynechococcus vulcanus* PsbO proteins. This alignment is similar to that presented by Popelkova and Yocum (2002). Binding Determinant I (highlighted in cyan) is found in all oxygenic organisms while Binding Determinant II (highlighted in magenta) is found only in the green plant lineage. **B.** Distance-constrained molecular dynamic refinements for spinach PsbO. Shown are the top ten models of PsbO presented in different shades of blue. These were aligned with cyanobacterial PsbO (orange, PDB: 3WU2). Binding Determinants I and II are outlined in cyan and magenta dashes, respectively. All of the models illustrated fulfill the EDC distance constraints and exhibit similar low DOPE scores (≈ -19,500). The orange arrow indicates the location of the cyano-loop (*T. vulcanus*, ^{131}N-^{138}S) while the blue arrow indicates the location of the short deletion (spinach, ^{178}R-^{180}D) in the cyanobacterial protein.

Endnotes

The FIGURES 4.1, 4.3, 4.4, 4.8 and TABLE 4.1 are reprinted with permission from the published article (cited below with the name of the author of this thesis underlined).

Manjula P. Mummadisetti, Laurie K. Frankel, Henry D. Bellamy, Larry Sallans, Jost S. Goettert, Michal Brylinski, and Terry M. Bricker (2016). Use of Protein Cross-Linking and Radiolytic Labeling To Elucidate the Structure of PsbO within Higher-Plant Photosystem II. Biochemistry, 55 (23), pp 3204–3213.

References

Acharya, K.R., and Lloyd, M.D. (2005). The advantages and limitations of protein crystal structures. Trends Pharmacol Sci *26*, 10-14.

Ago, H., Adachi, H., Umena, Y., Tashiro, T., Kawakami, K., Kamiya, N., Tian, L., Han, G., Kuang, T., Liu, Z., *et al.* (2016). Novel features of eukaryotic photosystem II revealed by its crystal structure analysis from a red alga. The Journal of Biological Chemistry.

Allahverdiyeva, Y., Suorsa, M., Rossi, F., Pavesi, A., Kater, M.M., Antonacci, A., Tadini, L., Pribil, M., Schneider, A., Wanner, G., *et al.* (2013). Arabidopsis plants lacking PsbQ and PsbR subunits of the oxygen-evolving complex show altered PSII super-complex organization and short-term adaptive mechanisms. The Plant Journal : for cell and molecular biology *75*, 671-684.

Aoi, M., Kashino, Y., and Ifuku, K. (2014). Function and association of CyanoP in photosystem II of Synechocystis sp PCC 6803. Res Chem Intermediates *40*, 3209-3217.

Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics *22*, 195-201.

Arnon, D.I. (1949). Copper enzymes in isolated chloroplasts. polyphenoloxidase in *Beta vulgaris*. Plant Physiology *24*, 1-15.

Balsera, M., Arellano, J.B., Revuelta, J.L., de las Rivas, J., and Hermoso, J.A. (2005). The 1.49 A resolution crystal structure of PsbQ from photosystem II of Spinacia oleracea reveals a PPII structure in the N-terminal region. Journal of Molecular Biology *350*, 1051-1060.

Banerjee, S., and Mazumdar, S. (2012). Electrospray ionization mass spectrometry: a technique to access the information beyond the molecular weight of the analyte. International Journal of Analytical Chemistry *2012*, 282574.

Boekema, E.J., van Breemen, J.F., van Roon, H., and Dekker, J.P. (2000). Conformational changes in photosystem II supercomplexes upon removal of extrinsic subunits. Biochemistry *39*, 12907-12915.
Bondar, A.N., and Dau, H. (2012). Extended protein/water H-bond networks in photosynthetic water oxidation. Biochimica et Biophysica Acta *1817*, 1177-1190.

Bricker, T., and Burnap, R. (2005). The Extrinsic Proteins of Photosystem II. In Photosystem II, T. Wydrzynski, K. Satoh, and J. Freeman, eds. (Springer Netherlands), pp. 95-120.

Bricker, T.M. (1992). Oxygen evolution in the absence of the 33-kilodalton manganese-stabilizing protein. Biochemistry *31*, 4623-4628.

Bricker, T.M., and Frankel, L.K. (1998). The structure and function of the 33 kDa extrinsic protein of Photosystem II: A critical assessment. Photosynthesis Research *56*, 157-173.

Bricker, T.M., and Frankel, L.K. (2011). Auxiliary functions of the PsbO, PsbP and PsbQ proteins of higher plant Photosystem II: a critical analysis. Journal of Photochemistry and Photobiology. B, Biology *104*, 165-178.

Bricker, T.M., Mummadisetti, M.P., and Frankel, L.K. (2015). Recent advances in the use of mass spectrometry to examine structure/function relationships in photosystem II. Journal of Photochemistry and Photobiology. B, Biology *152*, 227-246.

Bricker, T.M., Roose, J.L., Fagerlund, R.D., Frankel, L.K., and Eaton-Rye, J.J. (2012). The extrinsic proteins of Photosystem II. Biochimica et Biophysica Acta *1817*, 121-142.

Bricker, T.M., Roose, J.L., Zhang, P., and Frankel, L.K. (2013). The PsbP family of proteins. Photosynthesis Research *116*, 235-250.

Brzezowski, P., Wilson, K.E., and Gray, G.R. (2012). The PSBP2 protein of Chlamydomonas reinhardtii is required for singlet oxygen-dependent signaling. Planta *236*, 1289-1303.

Burnap, R.L., Shen, J.R., Jursinic, P.A., Inoue, Y., and Sherman, L.A. (1992). Oxygen yield and thermoluminescence characteristics of a cyanobacterium lacking the manganese-stabilizing protein of photosystem II. Biochemistry *31*, 7404-7410.

Burnap, R.L., and Sherman, L.A. (1991). Deletion mutagenesis in Synechocystis sp. PCC6803 indicates that the Mn-stabilizing protein of photosystem II is not essential for O2 evolution. Biochemistry *30*, 440-446.
Cao, P., Xie, Y., Li, M., Pan, X., Zhang, H., Zhao, X., Su, X., Cheng, T., and Chang, W. (2015). Crystal structure analysis of extrinsic PsbP protein of photosystem II reveals a manganese-induced conformational change. Molecular Plant 8, 664-666.

Cardona, T., Sedoud, A., Cox, N., and Rutherford, A.W. (2012). Charge separation in photosystem II: a comparative and evolutionary overview. Biochimica et Biophysica Acta *1817*, 26-43.

Cormann, K.U., Bartsch, M., Rogner, M., and Nowaczyk, M.M. (2014). Localization of the CyanoP binding site on photosystem II by surface plasmon resonance spectroscopy. Frontiers in Plant Science *5*, 595.
Darnell, S.J., LeGault, L., and Mitchell, J.C. (2008). KFC Server: interactive forecasting of protein interaction hot spots. Nucleic Acids Research *36*, W265-269.

De Las Rivas, J., and Barber, J. (2004). Analysis of the Structure of the PsbO Protein and its Implications. Photosynthesis Research *81*, 329-343.

de Vitry, C., Olive, J., Drapier, D., Recouvreur, M., and Wollman, F.A. (1989). Posttranslational events leading to the assembly of photosystem II protein complex: a study using photosynthesis mutants from Chlamydomonas reinhardtii. J Cell Biol. *109*, 991-1006.

Debus, R.J., Barry, B.A., Babcock, G.T., and McIntosh, L. (1988). Site-directed mutagenesis identifies a tyrosine radical involved in the photosynthetic oxygen-evolving system. Proceedings of the National Academy of Sciences of the United States of America *85*, 427-430.

Delepelaire, P., and Chua, N.H. (1979). Lithium Dodecyl Sulfate-Polyacrylamide Gel-Electrophoresis of Thylakoid Membranes at 4-Degrees-C - Characterizations of 2 Additional Chlorophyll - Protein Complexes. Proceedings of the National Academy of Sciences of the United States of America *76*, 111-115.

Demetrios F. Ghanotakis, G.T.B.a.C.F.Y. (1985). On the role of water-soluble polypeptides (17,23 kDa), calcium and chloride in photosynthetic oxygen evolution. The FEBS Journal *192*, 1-3.

El-Aneed, A., Cohen, A., and Banoub, J. (2009). Mass Spectrometry, Review of the Basics: Electrospray, MALDI, and Commonly Used Mass Analyzers. Appl Spectrosc Rev *44*, 210-230.

Eng, J.K., McCormack, A.L., and Yates, J.R. (1994). An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. Journal of the American Society for Mass Spectrometry *5*, 976-989.

Eswar, N., Webb, B., Marti-Renom, M.A., Madhusudhan, M.S., Eramian, D., Shen, M.Y., Pieper, U., and Sali, A. (2007). Comparative protein structure modeling using MODELLER. Current protocols in protein science / editorial board, John E. Coligan ... [et al.] *Chapter 2*, Unit 2 9.

Ferreira, K.N., Iverson, T.M., Maghlaoui, K., Barber, J., and Iwata, S. (2004). Architecture of the photosynthetic oxygen-evolving center. Science *303*, 1831-1838.

Forbush, B., Kok, B., McGloin, M. (1971). Cooperation of charges in photosynthetic O2 evolution-II. Damping of flash yield oscillation, deactivation. Photochemistry and Photobiology *14*, 307-321.

Frankel, L.K., Sallans, L., Limbach, P.A., and Bricker, T.M. (2012). Identification of oxidized amino acid residues in the vicinity of the Mn(4)CaO(5) cluster of Photosystem II: implications for the identification of oxygen channels within the Photosystem. Biochemistry *51*, 6371-6377.

Frankel, L.K., Sallans, L., Limbach, P.A., and Bricker, T.M. (2013). Oxidized amino acid residues in the vicinity of Q(A) and Pheo(D1) of the photosystem II reaction center: putative generation sites of reducing-side reactive oxygen species. PloS One *8*, e58042.

Gabdulkhakov, A., Guskov, A., Broser, M., Kern, J., Muh, F., Saenger, W., and Zouni, A. (2009). Probing the accessibility of the Mn(4)Ca cluster in photosystem II: channels calculation, noble gas derivatization, and cocrystallization with DMSO. Structure *17*, 1223-1234.

Gallivan, J.P., and Dougherty, D.A. (1999). Cation-pi interactions in structural biology. Proceedings of the National Academy of Sciences of the United States of America *96*, 9459-9464.

Geer, L.Y., Markey, S.P., Kowalak, J.A., Wagner, L., Xu, M., Maynard, D.M., Yang, X., Shi, W., and Bryant, S.H. (2004). Open mass spectrometry search algorithm. Journal of Proteome Research *3*, 958-964.

Ghanotakis, D.F., Babcock, G.T., and Yocum, C.F. (1984a). Calcium reconstitutes high rates of oxygen evolution in polypeptide depleted Photosystem II preparations. FEBS Letters *167*, 127-130.

Ghanotakis, D.F., Topper, J.N., Babcock, G.T., and Yocum, C.F. (1984b). Water-Soluble 17-Kda and 23-Kda Polypeptides Restore Oxygen Evolution Activity by Creating a High-Affinity Binding-Site for Ca-2+ on the Oxidizing Side of Photosystem-Ii. FEBS Letters *170*, 169-173.
Ghanotakis, K.K.D.F. (1991). Effect of the manganese complex on the binding of the extrinsic proteins (17, 23 and 33 kDa) of Photosystem II Photosynthesis Research *29*, 149-155.

Gotze, M., Pettelkau, J., Schaks, S., Bosse, K., Ihling, C.H., Krauth, F., Fritzsche, R., Kuhn, U., and Sinz, A. (2012). StavroX--a software for analyzing crosslinked products in protein interaction studies. Journal of the American Society for Mass Spectrometry *23*, 76-87.

Grabarek, Z., and Gergely, J. (1990). Zero-length crosslinking procedure with the use of active esters. Analytical Biochemistry *185*, 131-135.

GT, H. (1996). Bioconjugate Techniques (Academic, San Diego, CA).

Guskov, A., Kern, J., Gabdulkhakov, A., Broser, M., Zouni, A., and Saenger, W. (2009). Cyanobacterial photosystem II at 2.9-A resolution and the role of quinones, lipids, channels and chloride. Nature Structural & Molecular Biology *16*, 334-342.

Hall, M., Mata-Cabana, A., Akerlund, H.E., Florencio, F.J., Schroder, W.P., Lindahl, M., and Kieselbach, T. (2010). Thioredoxin targets of the plant chloroplast lumen and their implications for plastid function. Proteomics *10*, 987-1001.

Han, G., Mamedov, F., and Styring, S. (2012). Misses during water oxidation in photosystem II are S state-dependent. The Journal of Biological Chemistry *287*, 13422-13429.

Hansson, M., and Vener, A.V. (2003). Identification of three previously unknown in vivo protein phosphorylation sites in thylakoid membranes of Arabidopsis thaliana. Molecular & Cellular Proteomics : MCP *2*, 550-559.

Ho, F.M. (2012). Structural and mechanistic investigations of photosystem II through computational methods. Biochimica et Biophysica Acta *1817*, 106-120.

Ho, F.M., and Styring, S. (2008). Access channels and methanol binding site to the CaMn4 cluster in Photosystem II based on solvent accessibility simulations, with implications for substrate water access. Biochimica et Biophysica Acta *1777*, 140-153.

Ido, K., Ifuku, K., Yamamoto, Y., Ishihara, S., Murakami, A., Takabe, K., Miyake, C., and Sato, F. (2009). Knockdown of the PsbP protein does not prevent assembly of the dimeric PSII core complex but impairs accumulation of photosystem II supercomplexes in tobacco. Biochimica et Biophysica Acta *1787*, 873-881.
Ido, K., Kakiuchi, S., Uno, C., Nishimura, T., Fukao, Y., Noguchi, T., Sato, F., and Ifuku, K. (2012). The conserved His-144 in the PsbP protein is important for the interaction between the PsbP N-terminus and the Cyt b559 subunit of photosystem II. The Journal of Biological Chemistry *287*, 26377-26387.

Ido, K., Nield, J., Fukao, Y., Nishimura, T., Sato, F., and Ifuku, K. (2014). Cross-linking evidence for multiple interactions of the PsbP and PsbQ proteins in a higher plant photosystem II supercomplex. The Journal of Biological Chemistry *289*, 20150-20157.

Ifuku, K. (2015). Localization and functional characterization of the extrinsic subunits of photosystem II: an update. Bioscience, Biotechnology, and Biochemistry *79*, 1223-1231.

Ifuku, K., Ishihara, S., and Sato, F. (2010). Molecular functions of oxygen-evolving complex family proteins in photosynthetic electron flow. Journal of Integrative Plant Biology *52*, 723-734.

Ifuku, K., Ishihara, S., Shimamoto, R., Ido, K., and Sato, F. (2008). Structure, function, and evolution of the PsbP protein family in higher plants. Photosynthesis Research *98*, 427-437.

Ifuku, K., Nakatsu, T., Kato, H., and Sato, F. (2004). Crystal structure of the PsbP protein of photosystem II from Nicotiana tabacum. EMBO Reports *5*, 362-367.

Ifuku, K., Nakatsu, T., Shimamoto, R., Yamamoto, Y., Ishihara, S., Kato, H., and Sato, F. (2005a). Structure and function of the PsbP protein of photosystem II from higher plants. Photosynthesis Research *84*, 251-255.

Ifuku, K., and Sato, F. (2001). Importance of the N-terminal sequence of the extrinsic 23 kDa polypeptide in Photosystem II in ion retention in oxygen evolution. Biochimica et Biophysica Acta *1546*, 196-204.

Ifuku, K., and Sato, F. (2002). A truncated mutant of the extrinsic 23-kDa protein that absolutely requires the extrinsic 17-kDa protein for Ca2+ retention in photosystem II. Plant & Cell Physiology *43*, 1244-1249.

Ifuku, K., Yamamoto, Y., Ono, T., Ishihara, S., and Sato, F. (2005b). PsbP Protein, But Not PsbQ Protein, Is Essential for the Regulation and Stabilization of Photosystem II in Higher Plants. Plant Physiology *139*, 1175-1184.

Imre Vass, K.M.C., Zsuzsanna Deák, Steve R. mayes, James Barber (1992). Thermoluminescence and flash-oxygen characterization of the IC2 deletion mutant of Synechocystis sp. PCC 6803 lacking the Photosystem II 33 kDa protein. Biochim Biophys Acta-Bioenergetics *1102*, 195-201.

Ishihara, S., Takabayashi, A., Ido, K., Endo, T., Ifuku, K., and Sato, F. (2007). Distinct functions for the two PsbP-like proteins PPL1 and PPL2 in the chloroplast thylakoid lumen of Arabidopsis. Plant Physiol *145*, 668-679.

Ishikawa, Y., Schroder, W.P., and Funk, C. (2005). Functional analysis of the PsbP-like protein (sll1418) in Synechocystis sp. PCC 6803. Photosynthesis Research *84*, 257-262.

Jackson, S.A., and Eaton-Rye, J.J. (2015). Characterization of a Synechocystis sp. PCC 6803 double mutant lacking the CyanoP and Ycf48 proteins of Photosystem II. Photosynthesis Research *124*, 217-229.

Jackson, S.A., Hinds, M.G., and Eaton-Rye, J.J. (2012). Solution structure of CyanoP from Synechocystis sp. PCC 6803: new insights on the structural basis for functional specialization amongst PsbP family proteins. Biochimica et Biophysica Acta *1817*, 1331-1338.

Joliot, P. (1965). [Kinetics of photosynthetic fluorescence in relation to oxygen evolution]. Biochimica et Biophysica Acta *102*, 135-148.

Joliot, P. (1968). Kinetic studies of photosystem II in photosynthesis. Photochemistry and Photobiology *8*, 451-463.

Joliot, P. (2003). Period-four oscillations of the flash-induced oxygen formation in photosynthesis. Photosynthesis Research *76*, 65-72.

Joliot, P., Barbieri, G., Chabaud, R. (1969). A new model for the photochemical centers of the system II. Photosynthesis Research *10*, 309-329.

Kakiuchi, S., Uno, C., Ido, K., Nishimura, T., Noguchi, T., Ifuku, K., and Sato, F. (2012). The PsbQ protein stabilizes the functional binding of the PsbP protein to photosystem II in higher plants. Biochimica et Biophysica Acta *1817*, 1346-1351.

Kalkhof, S., and Sinz, A. (2008a). Chances and pitfalls of chemical cross-linking with amine-reactive N-hydroxysuccinimide esters. Analytical and Bioanalytical Chemistry *392*, 305-312.

Kalkhof, S., and Sinz, A. (2008b). Chances and pitfalls of chemical cross-linking with amine-reactive N-hydroxysuccinimide esters. Analytical and Bioanalytical Chemistry *392*, 305-312.

Kamiya, N., and Shen, J.R. (2003). Crystal structure of oxygen-evolving photosystem II from Thermosynechococcus vulcanus at 3.7-A resolution. Proceedings of the National Academy of Sciences of the United States of America *100*, 98-103.

Kashino, Y., Inoue-Kashino, N., Roose, J.L., and Pakrasi, H.B. (2006). Absence of the PsbQ protein results in destabilization of the PsbV protein and decreased oxygen evolution activity in cyanobacterial photosystem II. The Journal of Biological Chemistry *281*, 20834-20841.

Kaur, P., Kiselar, J.G., and Chance, M.R. (2009). Integrated algorithms for high-throughput examination of covalently labeled biomolecules by structural mass spectrometry. Analytical Chemistry *81*, 8141-8149.

Kern, J., Loll, B., Luneberg, C., DiFiore, D., Biesiadka, J., Irrgang, K.D., and Zouni, A. (2005). Purification, characterisation and crystallisation of photosystem II from Thermosynechococcus elongatus cultivated in a new type of photobioreactor. Biochimica et Biophysica Acta *1706*, 147-157.

Klockenbusch, C., O'Hara, J.E., and Kast, J. (2012). Advancing formaldehyde cross-linking towards quantitative proteomic applications. Analytical and Bioanalytical Chemistry *404*, 1057-1067.

Kok, B., Forbush, B., and McGloin, M. (1970). Cooperation of charges in photosynthetic O2 evolution-I. A linear four step mechanism. Photochemistry and Photobiology *11*, 457-475.

Kopecky, V., Jr., Kohoutova, J., Lapkouski, M., Hofbauerova, K., Sovova, Z., Ettrichova, O., Gonzalez-Perez, S., Dulebo, A., Kaftan, D., Smatanova, I.K., *et al.* (2012). Raman spectroscopy adds complementary detail to the high-resolution x-ray crystal structure of photosynthetic PsbP from Spinacia oleracea. PloS One *7*, e46694.

Kosinski, J., von Appen, A., Ori, A., Karius, K., Muller, C.W., and Beck, M. (2015). Xlink Analyzer: software for analysis and visualization of cross-linking data in the context of three-dimensional structures. Journal of Structural Biology *189*, 177-183.

Krewald, V., Retegan, M., Cox, N., Messinger, J., Lubitz, W., S., D., Neesea, F., and Pantazis, D.A. (2015). Metal oxidation states in biological water splitting. Chemical Science *6*, 1676-1695.

Kuhl, H., Rogner, M., Van Breemen, J.F., and Boekema, E.J. (1999). Localization of cyanobacterial photosystem II donor-side subunits by electron microscopy and the supramolecular organization ofphotosystem II in the thylakoid membrane. European Journal of Biochemistry / FEBS *266*, 453-459.

Kunkel, G.R., Mehrabian, M., and Martinson, H.G. (1981). Contact-site cross-linking agents. Molecular and Cellular Biochemistry *34*, 3-13.

Kurowski, M.A., and Bujnicki, J.M. (2003). GeneSilico protein structure prediction meta-server. Nucleic Acids Research *31*, 3305-3307.

Kuwabara, T., Murata, T., Miyao, M., and Murata, N. (1986). Partial Degradation of the 18-Kda Protein of the Photosynthetic Oxygen-Evolving Complex - a Study of a Binding-Site. Biochimica et Biophysica Acta *850*, 146-155.

Leuschner, C., and Bricker, T.M. (1996). Interaction of the 33 kDa extrinsic protein with photosystem II: rebinding of the 33 kDa extrinsic protein to photosystem II membranes which contain four, two, or zero manganese per photosystem II reaction center. Biochemistry *35*, 4551-4557.

Lichtenthaler, H.K. (1987). Chlorophylls and Carotenoids - Pigments of Photosynthetic Biomembranes. Method Enzymol *148*, 350-382.

Liu, H., Weisz, D.A., and Pakrasi, H.B. (2015). Multiple copies of the PsbQ protein in a cyanobacterial photosystem II assembly intermediate complex. Photosynthesis Research *126*, 375-383.

Liu, H., Zhang, H., Weisza, D.A., Vidavsky, I., Gross, M.L., and Pakrasi, H.B. (2014). MS-based cross-linking analysis reveals the location of the PsbQ protein in cyanobacterial photosystem II. Proceedings of the National Academy of Sciences of the United States of America *111*, 4638-4643.

Liu, J., Yang, H., Lu, Q., Wen, X., Chen, F., Peng, L., Zhang, L., and Lu, C. (2012). PsbP-domain protein1, a nuclear-encoded thylakoid lumenal protein, is essential for photosystem I assembly in Arabidopsis. The Plant Cell *24*, 4992-5006.

Liu., H., Frankel, L.K., and Bricker, T.M. (2007). Functional analysis of photosystem II in a PsbO-1-deficient mutant in Arabidopsis thaliana. Biochemistry *46*, 7607-7613.

Ljungberg, U., Jansson, C., Andersson, B., and Akerlund, H.E. (1983). Reconstitution of oxygen evolution in high salt washed photosystem II particles. Biochemical and Biophysical Research Communications *113*, 738-744.

Loll, B., Kern, J., Saenger, W., Zouni, A., and Biesiadka, J. (2005). Towards complete cofactor arrangement in the 3.0 A resolution structure of photosystem II. Nature *438*, 1040-1044.

Lorch, S., Capponi, S., Pieront, F., and Bondar, A.N. (2015). Dynamic Carboxylate/Water Networks on the Surface of the PsbO Subunit of Photosystem II. The Journal of Physical Chemistry. B *119*, 12172-12181.

Lovell, S.C., Davis, I.W., Arendall, W.B., 3rd, de Bakker, P.I., Word, J.M., Prisant, M.G., Richardson, J.S., and Richardson, D.C. (2003). Structure validation by Calpha geometry: phi,psi and Cbeta deviation. Proteins *50*, 437-450.

Luo, H., Jackson, S.A., Fagerlund, R.D., Summerfield, T.C., and Eaton-Rye, J.J. (2014). The importance of the hydrophilic region of PsbL for the plastoquinone electron acceptor complex of Photosystem II. Biochimica et Biophysica Acta *1837*, 1435-1446.

Madler, S., Bich, C., Touboul, D., and Zenobi, R. (2009). Chemical cross-linking with NHS esters: a systematic study on amino acid reactivities. Journal of Mass Spectrometry *44*, 694-706.

Mavankal, G., Mccain, D.C., and Bricker, T.M. (1986). Effects of Chloride on Paramagnetic Coupling of Manganese in Calcium Chloride-Washed Photosystem-Ii Preparations. FEBS Letters *202*, 235-239.

Mayfield, S.P., Bennoun, P., and Rochaix, J.D. (1987a). Expression of the nuclear encoded OEE1 protein is required for oxygen evolution and stability of photosystem II particles in Chlamydomonas reinhardtii. The EMBO Journal *6*, 313-318.

Mayfield, S.P., Rahire, M., Frank, G., Zuber, H., and Rochaix, J.D. (1987b). Expression of the nuclear gene encoding oxygen-evolving enhancer protein 2 is required for high levels of photosynthetic oxygen evolution

in Chlamydomonas reinhardtii. Proceedings of the National Academy of Sciences of the United States of America *84*, 749-753.

McEvoy, J.P., and Brudvig, G.W. (2006). Water-splitting chemistry of photosystem II. Chem Rev *106*, 4455-4483.

Meades, G.D., Jr., McLachlan, A., Sallans, L., Limbach, P.A., Frankel, L.K., and Bricker, T.M. (2005). Association of the 17-kDa extrinsic protein with photosystem II in higher plants. Biochemistry *44*, 15216-15221.

Michoux, F., Boehm, M., Bialek, W., Takasaka, K., Maghlaoui, K., Barber, J., Murray, J.W., and Nixon, P.J. (2014). Crystal structure of CyanoQ from the thermophilic cyanobacterium Thermosynechococcus elongatus and detection in isolated photosystem II complexes. Photosynthesis Research *122*, 57-67.

Michoux, F., Takasaka, K., Boehm, M., Nixon, P.J., and Murray, J.W. (2010). Structure of CyanoP at 2.8 A: implications for the evolution and function of the PsbP subunit of photosystem II. Biochemistry *49*, 7411-7413.

Migneault, I., Dartiguenave, C., Bertrand, M.J., and Waldron, K.C. (2004). Glutaraldehyde: behavior in aqueous solution, reaction with proteins, and application to enzyme crosslinking. BioTechniques *37*, 790-796, 798-802.

Mintseris, J., Pierce, B., Wiehe, K., Anderson, R., Chen, R., and Weng, Z. (2007). Integrating statistical pair potentials into protein complex prediction. Proteins *69*, 511-520.

Mitsue Miyao, N.M. (1989). The mode of binding of three extrinsic proteins of 33 kDa, 23 kDa and 18 kDa in the photosystem II complex of spinach. biochim Biophys Acta-Bioenergetics *3*, 315-321.

Miyao, M., Fujimura, Y., and Murata, N. (1988). Partial Degradation of the Extrinsic 23-Kda Protein of the Photosystem-Ii Complex of Spinach. Biochimica et Biophysica Acta *936*, 465-474.

Miyao, M., and Murata, N. (1983). Partial disintegration and reconstitution of the photosynthetic oxygen evolution system. Binding of 24 kilodalton and 18 kilodalton polypeptides. Biochimica et Biophysica Acta (BBA) - Bioenergetics *725*, 87-93.

Miyao, M., and Murata, N. (1984). Calcium ions can be substituted for the 24-kDa polypeptide in photosynthetic oxygen evolution. FEBS Letters *168*, 118-120.

Miyao, M., and Murata, N. (1985). The Cl− effect on photosynthetic oxygen evolution: interaction of Cl− with 18-kDa, 24-kDa and 33-kDa proteins. FEBS Letters *180*, 303-308.

Miyao, M., Murata, N. (1984). Role of the 33- kDa polypeptide in preserving Mn in the photosynthetic oxygen- evolution system and its replacement by chloride ions. FEBS Letters *170*, 350-354.

Miyao, M., Murata, N., Lavorel, J., Maison-Peteri, B., Boussac, A., Etienne, A. (1987). Effect of the 33-kDa protein on the S-state transitions in photosynthetic oxygen evolution. Biochimica et Biophysica Acta (BBA) - Bioenergetics *890*, 151-159.

Morris, A.L., MacArthur, M.W., Hutchinson, E.G., and Thornton, J.M. (1992). Stereochemical quality of protein structure coordinates. Proteins *12*, 345-364.

Mullineaux, C.W. (1999). The thylakoid membranes of cyanobacteria: structure, dynamics and function. Aust J Plant Physiol *26*, 671-677.

Mummadisetti, M.P., Frankel, L.K., Bellamy, H.D., Sallans, L., Goettert, J.S., Brylinski, M., Limbach, P.A., and Bricker, T.M. (2014). Use of protein cross-linking and radiolytic footprinting to elucidate PsbP and PsbQ interactions within higher plant Photosystem II. Proceedings of the National Academy of Sciences of the United States of America *111*, 16178-16183.

Murakami, R., Ifuku, K., Takabayashi, A., Shikanai, T., Endo, T., and Sato, F. (2002). Characterization of an Arabidopsis thaliana mutant with impaired psbO, one of two genes encoding extrinsic 33-kDa proteins in photosystem II. FEBS Letters *523*, 138-142.

Murakami, R., Ifuku, K., Takabayashi, A., Shikanai, T., Endo, T., and Sato, F. (2005). Functional dissection of two Arabidopsis PsbO proteins: PsbO1 and PsbO2. The FEBS Journal *272*, 2165-2175.

Murray, J.W., and Barber, J. (2007). Structural characteristics of channels and pathways in photosystem II including the identification of an oxygen channel. Journal of Structural Biology *159*, 228-237.

Nagao, R., Suzuki, T., Okumura, A., Niikura, A., Iwai, M., Dohmae, N., Tomo, T., Shen, J.R., Ikeuchi, M., and Enami, I. (2010). Topological analysis of the extrinsic PsbO, PsbP and PsbQ proteins in a green algal PSII complex by cross-linking with a water-soluble carbodiimide. Plant & Cell Physiology *51*, 718-727.

Nakatani, H.Y. (1984). Photosynthetic oxygen evolution does not require the participation of polypeptides of 16 and 24 kilodaltons. Biochemical and Biophysical Research Communications *120*, 299-304.
Nishimura, T., Nagao, R., Noguchi, T., Nield, J., Sato, F., and Ifuku, K. (2016). The N-terminal sequence of the extrinsic PsbP protein modulates the redox potential of Cyt b559 in photosystem II. Scientific Reports *6*, 21490.

Offenbacher, A.R., Polander, B.C., and Barry, B.A. (2013). An intrinsically disordered photosystem II subunit, PsbO, provides a structural template and a sensor of the hydrogen-bonding network in photosynthetic water oxidation. The Journal of Biological Chemistry *288*, 29056-29068.

Okumura, A., Nagao, R., Suzuki, T., Yamagoe, S., Iwai, M., Nakazato, K., and Enami, I. (2008). A novel protein in Photosystem II of a diatom Chaetoceros gracilis is one of the extrinsic proteins located on lumenal side and directly associates with PSII core components. Biochimica et Biophysica Acta *1777*, 1545-1551.

Ono, T., Izawa, S., and Inoue, Y. (1992). Structural and functional modulation of the manganese cluster in Ca(2+)-depleted photosystem II induced by binding of the 24-kilodalton extrinsic protein. Biochemistry *31*, 7648-7655.

Pearson, W.R. (2013). An introduction to sequence similarity ("homology") searching. Current protocols in bioinformatics / editorial board, Andreas D. Baxevanis ... [et al.] *Chapter 3*, Unit3.1.

Peltier, J.B., Emanuelsson, O., Kalume, D.E., Ytterberg, J., Friso, G., Rudella, A., Liberles, D.A., Soderberg, L., Roepstorff, P., von Heijne, G., *et al.* (2002). Central functions of the lumenal and peripheral thylakoid proteome of Arabidopsis determined by experimentation and genome-wide prediction. The Plant Cell *14*, 211-236.

Peri, S., Steen, H., and Pandey, A. (2001). GPMAW--a software tool for analyzing proteins and peptides. Trends in Biochemical Sciences *26*, 687-689.

Perkins, D.N., Pappin, D.J., Creasy, D.M., and Cottrell, J.S. (1999). Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis *20*, 3551-3567.

Petrek, M., Otyepka, M., Banas, P., Kosinova, P., Koca, J., and Damborsky, J. (2006). CAVER: a new tool to explore routes from protein clefts, pockets and cavities. BMC Bioinformatics *7*, 316.

Phuthong, W., Huang, Z., Wittkopp, T.M., Sznee, K., Heinnickel, M.L., Dekker, J.P., Frese, R.N., Prinz, F.B., and Grossman, A.R. (2015). The Use of Contact Mode Atomic Force Microscopy in Aqueous Medium for Structural Analysis of Spinach Photosynthetic Complexes. Plant Physiology *169*, 1318-1332.

Popelka, H., and Yocum, C. (2012). Probing the N-terminal sequence of spinach PsbO: evidence that essential threonine residues bind to different functional sites in eukaryotic photosystem II. Photosynthesis Research *112*, 117-128.

Popelkova, H., Commet, A., Kuntzleman, T., and Yocum, C.F. (2008). Inorganic cofactor stabilization and retention: the unique functions of the two PsbO subunits of eukaryotic photosystem II. Biochemistry *47*, 12593-12600.

Popelkova, H., Im, M.M., D'Auria, J., Betts, S.D., Lydakis-Simantiris, N., and Yocum, C.F. (2002a). N-terminus of the photosystem II manganese stabilizing protein: effects of sequence elongation and truncation. Biochemistry *41*, 2702-2711.

Popelkova, H., Im, M.M., and Yocum, C.F. (2002b). N-terminal truncations of manganese stabilizing protein identify two amino acid sequences required for binding of the eukaryotic protein to photosystem II and reveal the absence of one binding-related sequence in cyanobacteria. Biochemistry *41*, 10038-10045.

Popelkova, H., Im, M.M., and Yocum, C.F. (2003). Binding of manganese stabilizing protein to photosystem II: identification of essential N-terminal threonine residues and domains that prevent nonspecific binding. Biochemistry *42*, 6193-6200.

Pospisil, P. (2009). Production of reactive oxygen species by photosystem II. Biochimica et Biophysica Acta *1787*, 1151-1160.

Pospisil, P. (2014). The role of metals in production and scavenging of reactive oxygen species in photosystem II. Plant & Cell Physiology *55*, 1224-1232.

Pospísil, P., Arató, A., Krieger-Liszkay, A., and Rutherford, A.W. (2004). Hydroxyl radical generation by photosystem II. Biochemistry *43*, 6783-6792.

Puthiyaveetil, S., and Kirchhoff, H. (2013). A phosphorylation map of the photosystem II supercomplex C2S2M2. Frontiers in Plant Science *4*, 459.

Rathner, P., Rathner, A., Hornicakova, M., Wohlschlager, C., Chandra, K., Kohoutova, J., Ettrich, R., Wimmer, R., and Muller, N. (2015). Solution NMR and molecular dynamics reveal a persistent alpha helix within the dynamic region of PsbQ from photosystem II of higher plants. Proteins *83*, 1677-1686.

Reiland, S., Messerli, G., Baerenfaller, K., Gerrits, B., Endler, A., Grossmann, J., Gruissem, W., and Baginsky, S. (2009). Large-scale Arabidopsis phosphoproteome profiling reveals novel chloroplast kinase substrates and phosphorylation networks. Plant Physiol *150*, 889-903.

Roose, J.L., Frankel, L.K., and Bricker, T.M. (2010). Documentation of significant electron transport defects on the reducing side of photosystem II upon removal of the PsbP and PsbQ extrinsic proteins. Biochemistry *49*, 36-41.

Roose, J.L., Frankel, L.K., and Bricker, T.M. (2011a). Developmental defects in mutants of the PsbP domain protein 5 in Arabidopsis thaliana. PloS One *6*, e28624.

Roose, J.L., Frankel, L.K., and Bricker, T.M. (2014). The PsbP domain protein 1 functions in the assembly of lumenal domains in photosystem I. The Journal of Biological Chemistry *289*, 23776-23785.

Roose, J.L., Frankel, L.K., Mummadisetti, M.P., and Bricker, T.M. (2016). The extrinsic proteins of photosystem II: update. Planta.

Roose, J.L., Kashino, Y., and Pakrasi, H.B. (2007a). The PsbQ protein defines cyanobacterial Photosystem II complexes with highest activity and stability. Proceedings of the National Academy of Sciences of the United States of America *104*, 2548-2553.
Roose, J.L., Wegener, K.M., and Pakrasi, H.B. (2007b). The extrinsic proteins of Photosystem II. Photosynthesis Research *92*, 369-387.

Roose, J.L., Yocum, C.F., and Popelkova, H. (2011b). Binding stoichiometry and affinity of the manganese-stabilizing protein affects redox reactions on the oxidizing side of photosystem II. Biochemistry *50*, 5988-5998.

Rozbesky, D., Man, P., Kavan, D., Chmelik, J., Cerny, J., Bezouska, K., and Novak, P. (2012). Chemical cross-linking and H/D exchange for fast refinement of protein crystal structure. Analytical Chemistry *84*, 867-870.

Saito, K., Rutherford, A.W., and Ishikita, H. (2013a). Mechanism of proton-coupled quinone reduction in Photosystem II. Proceedings of the National Academy of Sciences of the United States of America *110*, 954-959.

Saito, K., Rutherford, A.W., and Ishikita, H. (2013b). Mechanism of tyrosine D oxidation in Photosystem II. Proceedings of the National Academy of Sciences of the United States of America *110*, 7690-7695.

Sato, N. (2010). Phylogenomic and structural modeling analyses of the PsbP superfamily reveal multiple small segment additions in the evolution of photosystem II-associated PsbP protein in green plants. Molecular Phylogenetics and Evolution *56*, 176-186.

Schubert, M., Petersson, U.A., Haas, B.J., Funk, C., Schroder, W.P., and Kieselbach, T. (2002). Proteome map of the chloroplast lumen of Arabidopsis thaliana. The Journal of Biological Chemistry *277*, 8354-8365.
Scigelova, M., Hornshaw, M., Giannakopulos, A., and Makarov, A. (2011). Fourier transform mass spectrometry. Molecular & Cellular Proteomics : MCP *10*, M111 009431.

Sharma, J., Panico, M., Barber, J., and Morris, H.R. (1997). Purification and determination of intact molecular mass by electrospray ionization mass spectrometry of the photosystem II reaction center subunits. The Journal of Biological Chemistry *272*, 33153-33157.

Shen, J.R. (1997). Possible functional differences between dimer and monomer of Photosystem II complex (Kluwer Academic Publishers).

Shen, M.Y., and Sali, A. (2006). Statistical potential for assessment and prediction of protein structures. Protein Science : a publication of the Protein Society *15*, 2507-2524.

Shukla, A.K., and Futrell, J.H. (2000). Tandem mass spectrometry: dissociation of ions by collisional activation. Journal of Mass Spectrometry : JMS *35*, 1069-1090.

Shutova, T., Klimov, V.V., Andersson, B., and Samuelsson, G. (2007). A cluster of carboxylic groups in PsbO protein is involved in proton transfer from the water oxidizing complex of Photosystem II. Biochimica et Biophysica Acta *1767*, 434-440.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J., *et al.* (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular Systems Biology *7*, 539.

Sinz, A. (2006). Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions. Mass Spectrometry Reviews *25*, 663-682.

Sleno, L., and Volmer, D.A. (2004). Ion activation methods for tandem mass spectrometry. Journal of Mass Spectrometry : JMS *39*, 1091-1112.

Smith, P.K., Krohn, R.I., Hermanson, G.T., Mallia, A.K., Gartner, F.H., Provenzano, M.D., Fujimoto, E.K., Goeke, N.M., Olson, B.J., and Klenk, D.C. (1985). Measurement of Protein Using Bicinchoninic Acid. Analytical Biochemistry *150*, 76-85.

Sriswasdi, S., Harper, S.L., Tang, H.Y., and Speicher, D.W. (2014). Enhanced identification of zero-length chemical cross-links using label-free quantitation and high-resolution fragment ion spectra. Journal of Proteome Research *13*, 898-914.

Staros, J.V. (1982). Membrane-impermeant, cleavable cross-linkers: new probes of nearest neighbor relationships at one face of a membrane. Biophysical Journal *37*, 21-22.

Suga, M., Akita, F., Hirata, K., Ueno, G., Murakami, H., Nakajima, Y., Shimizu, T., Yamashita, K., Yamamoto, M., Ago, H., *et al.* (2015). Native structure of photosystem II at 1.95 A resolution viewed by femtosecond X-ray pulses. Nature *517*, 99-103.

Suorsa, M., Regel, R.E., Paakkarinen, V., Battchikova, N., Herrmann, R.G., and Aro, E.M. (2004). Protein assembly of photosystem II and accumulation of subcomplexes in the absence of low molecular mass subunits PsbL and PsbJ. European Journal of Biochemistry / FEBS *271*, 96-107.

Sutherland, B.W., Toews, J., and Kast, J. (2008). Utility of formaldehyde cross-linking and mass spectrometry in the study of protein-protein interactions. Journal of Mass Spectrometry : JMS *43*, 699-715.

Swaim, C.L., Smith, J.B., and Smith, D.L. (2004). Unexpected products from the reaction of the synthetic cross-linker 3,3'-dithiobis(sulfosuccinimidyl propionate), DTSSP with peptides. Journal of the American Society for Mass Spectrometry *15*, 736-749.

Taka-aki Ono, Y.I. (1985). S-state turnover in the O2-evolving system of CaCl2-washed Photosystem II particles depleted of three peripheral proteins as measured by thermoluminescence. Removal of 33 kDa protein inhibits S3 to S4 transition. BBA Bioenergetics *806*, 331-340.

Takamoto, K., and Chance, M.R. (2006). Radiolytic protein footprinting with mass spectrometry to probe the structure of macromolecular complexes. Annual Review of Biophysics and Biomolecular Structure *35*, 251-276.

Thornton, L.E., Ohkawa, H., Roose, J.L., Kashino, Y., Keren, N., and Pakrasi, H.B. (2004). Homologs of plant PsbP and PsbQ proteins are necessary for regulation of photosystem ii activity in the cyanobacterium Synechocystis 6803. The Plant Cell *16*, 2164-2175.

Tikkanen, M., and Aro, E.M. (2012). Thylakoid protein phosphorylation in dynamic regulation of photosystem II in higher plants. Biochimica et Biophysica Acta *1817*, 232-238.
Toews, J., Rogalski, J.C., Clark, T.J., and Kast, J. (2008). Mass spectrometric identification of formaldehyde-induced peptide modifications under in vivo protein cross-linking conditions. Analytica Chimica Acta *618*, 168-183.

Tomita, M., Ifuku, K., Sato, F., and Noguchi, T. (2009). FTIR evidence that the PsbP extrinsic protein induces protein conformational changes around the oxygen-evolving Mn cluster in photosystem II. Biochemistry *48*, 6318-6325.

Tovchigrechko, A., and Vakser, I.A. (2005). Development and testing of an automated approach to protein docking. Proteins *60*, 296-301.

Turkina, M.V., Kargul, J., Blanco-Rivero, A., Villarejo, A., Barber, J., and Vener, A.V. (2006). Environmentally modulated phosphoproteome of photosynthetic membranes in the green alga Chlamydomonas reinhardtii. Molecular & Cellular Proteomics : MCP *5*, 1412-1425.

Umena, Y., Kawakami, K., Shen, J.R., and Kamiya, N. (2011). Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 A. Nature *473*, 55-60.

Vassiliev, S., Comte, P., Mahboob, A., and Bruce, D. (2010). Tracking the flow of water through photosystem II using molecular dynamics and streamline tracing. Biochemistry *49*, 1873-1881.

Veerman, J., Bentley, F.K., Eaton-Rye, J.J., Mullineaux, C.W., Vasil'ev, S., and Bruce, D. (2005). The PsbU subunit of photosystem II stabilizes energy transfer and primary photochemistry in the phycobilisome-photosystem II assembly of Synechocystis sp. PCC 6803. Biochemistry *44*, 16939-16948.

Vener, A.V. (2007). Environmentally modulated phosphorylation and dynamics of proteins in photosynthetic membranes. Biochimica et Biophysica Acta *1767*, 449–457.

Vener, A.V., Harms, A., Sussman, M.R., and Vierstra, R.D. (2001). Mass spectrometric resolution of reversible protein phosphorylation in photosynthetic membranes of Arabidopsis thaliana. The Journal of Biological Chemistry *276*, 6959-6966.

Vyas, V.K., Ukawala, R.D., Ghate, M., and Chintha, C. (2012). Homology modeling a fast tool for drug discovery: current perspectives. Indian Journal of Pharmaceutical Sciences *74*, 1-17.

Watanabe, M., Iwa, M., Narikawa, R., and Ikeuchi, M. (2009). Is the Photosystem II Complex a Monomer or a Dimer? . Plant Cell Physiology *50*, 1674-1680.

Xu, H., and Freitas, M.A. (2007a). A mass accuracy sensitive probability based scoring algorithm for database searching of tandem mass spectrometry data. BMC Bioinformatics *8*.

Xu, H., and Freitas, M.A. (2007b). A mass accuracy sensitive probability based scoring algorithm for database searching of tandem mass spectrometry data. BMC Bioinformatics *8*, 133.

Xu, H., and Freitas, M.A. (2008). Monte carlo simulation-based algorithms for analysis of shotgun proteomic data. Journal of Proteome Research *7*, 2605-2615.

Xu, H., and Freitas, M.A. (2009). MassMatrix: a database search program for rapid characterization of proteins and peptides from tandem mass spectrometry data. Proteomics *9*, 1548-1555.

Yamamoto, Y., Nakayama, S., Cohn, C.L., and Krogmann, D.W. (1987). Highly efficient purification of the 33-, 24-, and 18-kDa proteins in spinach photosystem II by butanol/water phase partitioning and high-performance liquid chromatography. Archives of Biochemistry and Biophysics *255*, 156-161.

Yang, B., Wu, Y.J., Zhu, M., Fan, S.B., Lin, J., Zhang, K., Li, S., Chi, H., Li, Y.X., Chen, H.F., *et al.* (2012). Identification of cross-linked peptides from complex samples. Nature Methods *9*, 904-906.

Yano, J., Kern, J., Sauer, K., Latimer, M.J., Pushkar, Y., Biesiadka, J., Loll, B., Saenger, W., Messinger, J., Zouni, A., *et al.* (2006). Where water is oxidized to dioxygen: structure of the photosynthetic Mn4Ca cluster. Science *314*, 821-825.

Yee, A.A., Savchenko, A., Ignachenko, A., Lukin, J., Xu, X.H., Skarina, T., Evdokimova, E., Liu, C.S., Semesi, A., Guido, V., *et al.* (2005). NMR and x-ray crystallography, complementary tools in structural proteomics of small proteins. Journal of the American Chemical Society *127*, 16512-16517.

Yi, X., Hargett, S.R., Liu, H., Frankel, L.K., and Bricker, T.M. (2007). The PsbP protein is required for photosystem II complex assembly/stability and photoautotrophy in Arabidopsis thaliana. The Journal of Biological Chemistry *282*, 24833-24841.

Zouni, A., Witt, H.T., Kern, J., Fromme, P., Krauss, N., Saenger, W., and Orth, P. (2001). Crystal structure of photosystem II from Synechococcus elongatus at 3.8 A resolution. Nature *409*, 739-743.

Zubrzycki, I.Z., Frankel, L.K., Russo, P.S., and Bricker, T.M. (1998). Hydrodynamic studies on the manganese-stabilizing protein of photosystem II. Biochemistry *37*, 13553-13558.

CHAPTER 5.

SUMMARY AND CONCLUSIONS

LOCATIONS OF EXTRINSIC PROTEINS IN HIGHER PLANT PHOTOSYSTEM II

This dissertation examined the structural aspects of the proteins PsbO, PsbP and PsbQ from spinach PSII membranes using protein crosslinking coupled to high-resolution tandem mass spectrometry. These extrinsic proteins are essential in maintaining the overall conformation around the Mn_4CaO_5 cluster, optimizing the ionic environment (Ca^{+2} and Cl^-) at the site of water oxidation, protect this site from exogenous reductants, and stabilizing the PSII function. X-ray crystal structures of spinach PsbP and PsbQ are available at 1.9 Å (Cao et al., 2015) and 1.45 Å (Balsera et al., 2005) respectively, however, no structures are available for higher plant PsbO. Cyanobacterial PsbO, however, is present in the crystal structure (3ARC) from *T. vulcanus*. The cyanobacterial PsbO has 47% sequence identity and about 60% sequence similarity with spinach PsbO. These two PsbO's are similar in their overall fold, however they do have a number of structural differences (as shown in the threaded structure of higher plant PsbO compared to cyanobacterial PsbO, FIGURE 1.5). Additionally, cyanobacterial PSII structure contains one copy of PsbO per PSII monomer, however, biochemical studies suggest two copies of PsbO per PSII monomer in higher plants (refer section 1.33). In any case, PsbO structure and its location in cyanobacterial PSII is a good model to understand its higher plant forms. PsbP and PsbQ, on the other hand, are not present in cyanobacteria, and their crystal structures (from higher plant) do not represent their bound forms to PSII. The structural homologs of PsbP and PsbQ in cyanobacteria are called CyanoP and CyanoQ respectively, which are functionally different from their higher plant forms. The X-ray crystal structures of soluble CyanoP and CyanoQ are available at 2.8 Å and 1.8 Å resolution respectively. However, these proteins are absent in the current PSII crystal structures, and hence their binding location in PSII remains elusive.

The three-dimensional structure of a protein also affects its function. Identifying the location of these extrinsic proteins would allow us to understand their structure-function relationships. Protein crosslinking studies have become one of the alternate methods to capture the structural information not available by other methods. It can be useful in identifying the relative positions of the unresolved regions of proteins in a crystal structure, arrest an unstable intermediate, and hence, identify the interacting domains or interacting partners in a protein complex. Since location of the extrinsic proteins in higher plant PSII has been debated, this method has become valuable in determining the interacting regions of these proteins, allowing us to propose models helpful in understanding the overall architecture of extrinsic proteins in PSII complex. Such studies have been performed for CyanoQ (see section 1.4.1 for details) (Liu et al., 2014). The authors identified the putative domains of interaction between CyanoQ-PsbO, CyanoQ-CyanoQ and CyanoQ-CP47, indicating that CyanoQ is present in the vicinity of PsbO and CP47. These results allowed identification of the binding location of CyanoQ to cyanobacterial PSII. Similarly, the N-terminus of PsbP (^{1}A) was crosslinked to Cyt b_{559} (^{57}E) with the EDC crosslinker (Ido et al., 2014), which indicates that these two proteins are within the van der Waals distance to each other. Additionally, PsbP was also found to closely interact with PsbR and CP26.

Using a similar approach, this dissertation describes the results obtained from protein crosslinking coupled to tandem mass spectrometry in PSII. My work identified putative interacting domains within PsbP, PsbQ, PsbO and between PsbP-PsbQ. The residues from the N-terminus of PsbP were crosslinked to the C-terminus of PsbP with the crosslinker BS3, which suggested that the N-terminus closely associates with its C-terminal domain. The structure of the N-terminus has not been determined in any of the PsbP structures available till date. The hypothesis that the PsbP interacts with PSII with an extended N-terminus, is thus rejected, it instead has a compact structure

when bound to PSII. FTIR studies indicated conformational changes around the Mn_4CaO_5 cluster upon PsbP and PsbQ binding to PSII (Kakiuchi et al., 2012; Tomita et al., 2009). Protein crosslinking with BS3 identified two crosslinked residue-pairs (^{40}K – ^{155}K and ^{40}K – ^{157}K) in PsbP and one pair (^{125}K – ^{133}Y) in PsbQ which were >11.4 Å apart when mapped onto their soluble forms of crystal structures. The distances greater than the crosslinker's active radius could mean that a conformational change occurred within these proteins upon their binding to PSII. Additionally, the residues (^{93}Y, ^{96}K and ^{97}T) in the loop ^{90}K-^{107}V of PsbP were crosslinked to the N-terminal ^{1}E of PsbQ, indicating the interacting domain of PsbP with PsbQ in PSII. This is the first report of a direct interaction between these two proteins.

Additionally, the residues ^{98}K – ^{133}Y and ^{101}K - ^{133}Y of PsbQ were also crosslinked with BS3. The distances between these residues were >30 Å when mapped onto the PsbQ structure (1VYK), which would only be possible if two copies of PsbQ interact within the photosystem. Similar findings were reported for CyanoQ earlier in cyanobacteria (Liu et al., 2014), and the putative binding location for CyanoQ was presented. These authors identified the residue pairs CyanoQ: ^{120}K – PsbO: ^{59}K, CyanoQ: ^{120}K – PsbO: ^{180}K crosslinked with a 12 Å-crosslinker (DTSSP) and CyanoQ: ^{102}K – PsbB: ^{440}D with a zero-length crosslinker (EDC). These distance constraints were incorporated in their models shown in FIGURE 5.1 A. In this model (Liu et al., 2014) the Pakrasi group proposed the binding of two CyanoQ molecules in an anti-parallel fashion within a PSII dimer, where PsbQ was placed between PsbO and PsbB (CP47).

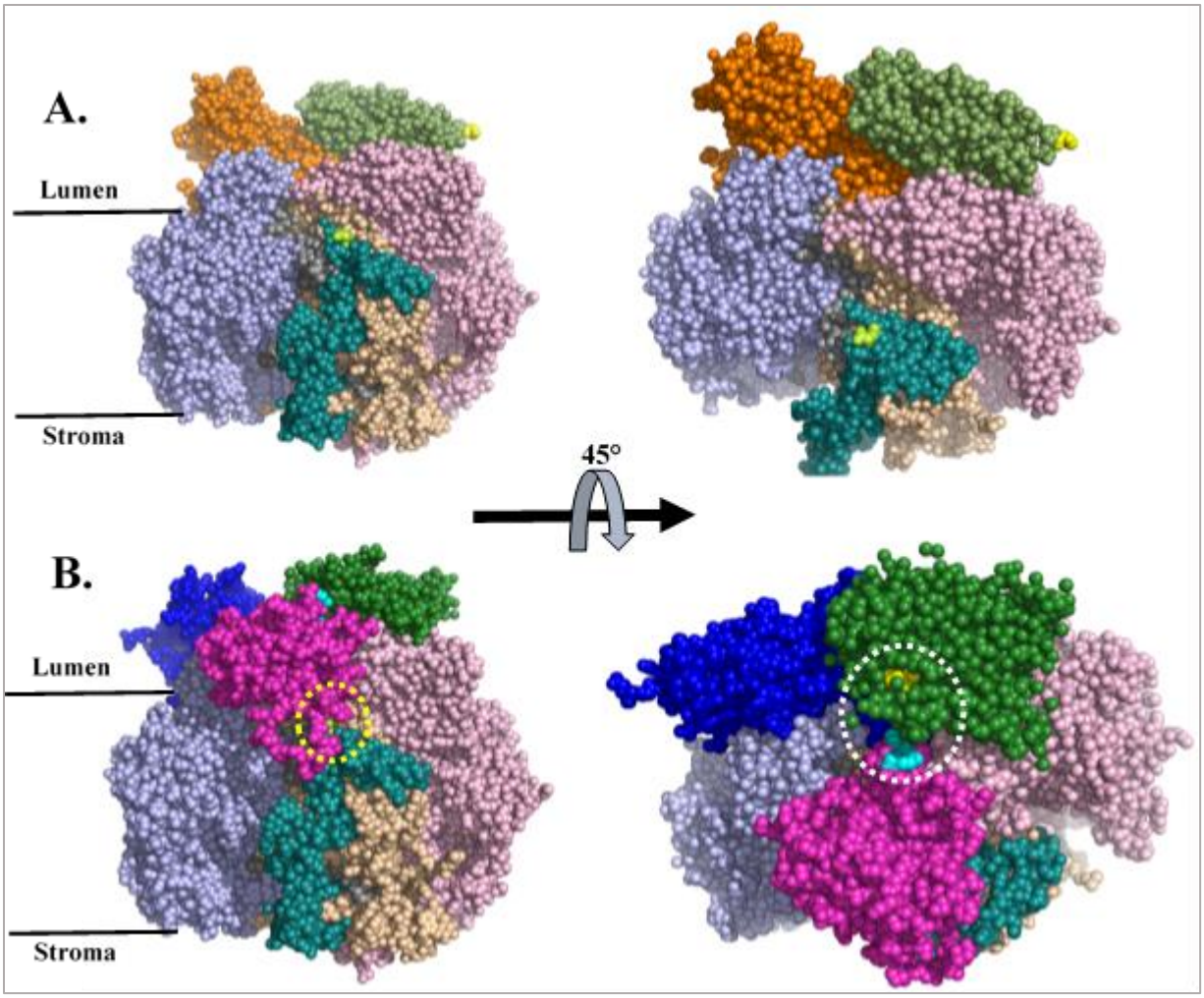


FIGURE 5.1. Models for the Location of Higher Plant Extrinsic Proteins.

The models presented in both **A.** and **B**. are modified from (Roose et al., 2016). These figures correspond to the structure of PSII, proposed by Pakrasi group (Liu et al., 2014). PsbE is shown in teal, PsbA in light grey, PsbB in light pink, PsbC in pale blue, PsbD in pale wheat. PsbO in orange (cyanobacteria) and blue (spinach), PsbQ in smudge green (cyanobacteria) and forest green (spinach), and PsbP in magenta.

A. Left, PSII core with proposed model of CyanoQ between PsbB and PsbO (Liu et al., 2014). Right, view of same model from lumen to stroma. **B.** Proposed model with PsbQ similar in location to CyanoQ in cyanobacteria. The P-Q interacting domain is represented with a dotted circle. The models generated in Chapter 3 and Chapter 4 are used for the representation here.

Based on their model, with the structural similarity and a resemblance in the crosslinking pattern between CyanoQ and PsbQ (inter-PsbQ interaction in an anti-parallel fashion), the location of higher plant PsbQ is proposed and illustrated in FIGURE 5.1 B. In this figure, location of PsbP is modelled based on our crosslinking data for PsbP-PsbQ interaction and from PsbP-PsbE interaction (Ido et al., 2012). Our models presented here are consistent with the crosslinker's distance constraints. These models are constrained to have ^{1}A of PsbP (shown as a cyan colored sphere) positioned within a van der Waals distance of PsbE: ^{57}E, and ^{93}Y, ^{96}K and ^{97}T residues (shown in yellow spheres) of PsbP around 11.4 Å from the N-terminus ^{1}E of PsbQ (not visible in these models). In Figure 5.1B, the CyanoQ structure is replaced with PsbQ and docked in a similar location as CyanoQ. Additionally, one of the lowest energy models for PsbP and PsbO, (presented in Chapter 3 and 4) were used while building these models. In this structure (Liu et al., 2014), the PsbQ is positioned between the two monomers in a PSII dimer.

Another possibility exists for the location of PsbQ. Instead of each PsbQ from a monomer interacting in a PSII dimer (Liu et al., 2014), the PsbQ-PsbQ crosslinking could be a result of interaction of PsbQ molecules between two different PSII-dimers. This model is consistent with the location of Q' seen in red-algal crystal structure (Ago et al., 2016). A model with PsbQ in the location of Q' is proposed in the FIGURE 5.2B. In first model (FIGURE 5.1), PsbQ is located closer to CP47, on the other hand, in the second model (FIGURE 5.2), the PsbQ is located near CP43. The figures shown in FIGURE 5.1 and 5.2, are modified from (Roose et al., 2016). The models presented here only contain core PSII protiens PsbA (D1), PsbB (CP47), PsbC (CP43), and PsbD (D2) along with the extrinsic proteins PsbO, PsbP, PsbQ, CyanoQ and PsbQ' (whichever is relavent). The N-terminus of PsbQ' is not completely resolved, the first resolved residue in red algae crystal structure for PsbQ' is Leu 109, and hence location of its N-terminus is not known.

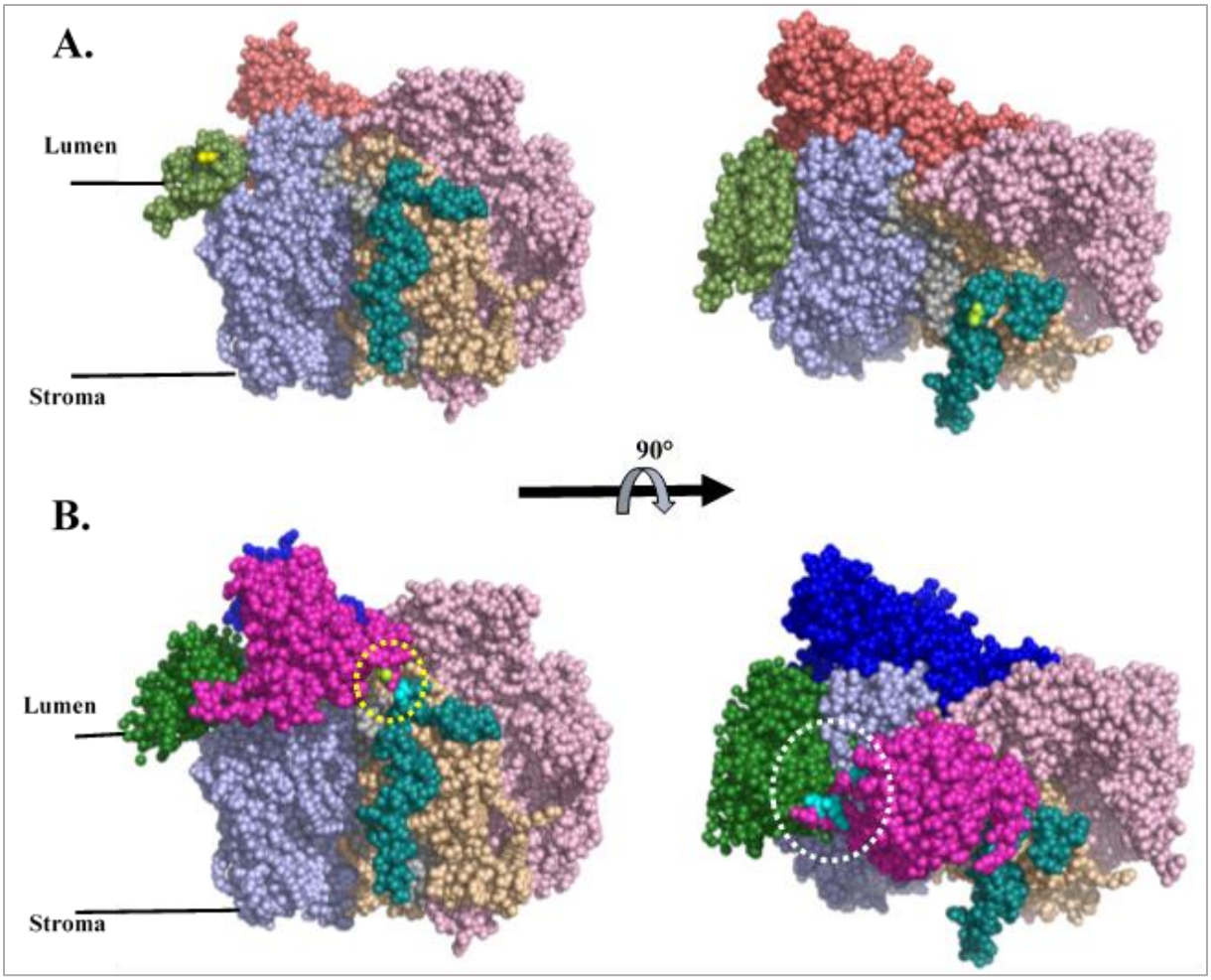


FIGURE 5.2. Models for the Location of PsbP and PsbQ Based on Protein-Crosslinking

The models presented in both A. and B. are modified from (Roose et al., 2016). **A.** Left, PSII core along with PsbO and PsbQ' in red algae crystal structure (Ago et al., 2016). The ^{109}L is the first residue of PsbQ resolved in this red algae structure, and are represented by yellow speheres. Right, 90° of red algal structure. This view is looking at PSII core from the lumen towards stroma. PsbE: ^{57}E is shown in lemon colored spheres **B.** Proposed location of PsbQ based on PsbQ'. PsbQ' is replaced with PsbQ. And red algal PsbO is replaced with spinach PsbO modelled from crosslkinking data. Left, PSII core with higher plant extrinsics, the N-terminal PsbP: ^{1}A is very close to PsbE: ^{59}E, shown with a yellow dotted circle. Right, 90° rotation of the model from left. This view is from lumen to stroma. The interacting residues between PsbP and PsbQ are highlighted with a white dotted circle. D1 (grey), CP47 (light pink), CP43 (light blue), D2 (wheat), Cytb*559* α-subunit (PsbE) in deapteal, PsbQ' (smudge green), PsbQ (forest green), PsbP (magenta), red algal PsbO (salmon), spinach PsbO (dark blue).

When PsbQ is placed in a similar location as PsbQ', it required much rearrangement for the placement of PsbP to satisfy the distance constraints from PsbQ and PsbE. Thus using crosslinking helped us in proposing models for the location of higher plant extrinsic proteins by incorporating the knowledge available for PSII.

Additionally, the EDC crosslinking results described in Chapter 4, for PsbO, allowed us to build models, which incorporated the distance constraints from the crosslinked-residues. EDC-crosslinked residues were constrained to have their Cα- Cα within 12 Å (see Chapter 4 for details). The spinach PsbO has a ten amino-acid extention at its N-terminus, which is absent in cyanobacterial PsbO. The crosslinking results allowed us to model this N-terminal domain, which facilitated in identifying the two binding determinants, required for the complete binding of two copies of PsbO. One of the binding determinant (^{4}K-^{10}E) is found only in higher plants (and green algae), but not found in cyanobacteria. The deletion of this domain reduces the binding ability of PsbO to 50%. The other determinant (^{15}T-^{18}E) is present in both higher plants and cyanobacteria. Deletion of this domain causes further loss of PsbO binding and consequently very poor oxygen evolving rates. Thus our model for PsbO enhanced our knowledge on the structure of PsbO from higher plants, and this model serves as a low-resolution structure. In FIGURE 5.1B and 5.2B, one of the lowest energy models of PsbO is used. Thus protein crosslinking provided considerable new information allowing building models pertaining to the location of extrinsic proteins in higher plant PSII.

Very recently, Wei and his coworkers identified the structure of higher plant PSII-supercomplex using membranes (Wei et al., 2016). This structure was resolved at 3.2 Å, using cryo-Electron microscopy. In this structure, the location of extrinsic proteins appeared to form a triangular crest on PSII membrane proteins when viewed from lumenal to stromal face. The extrinsic proteins,

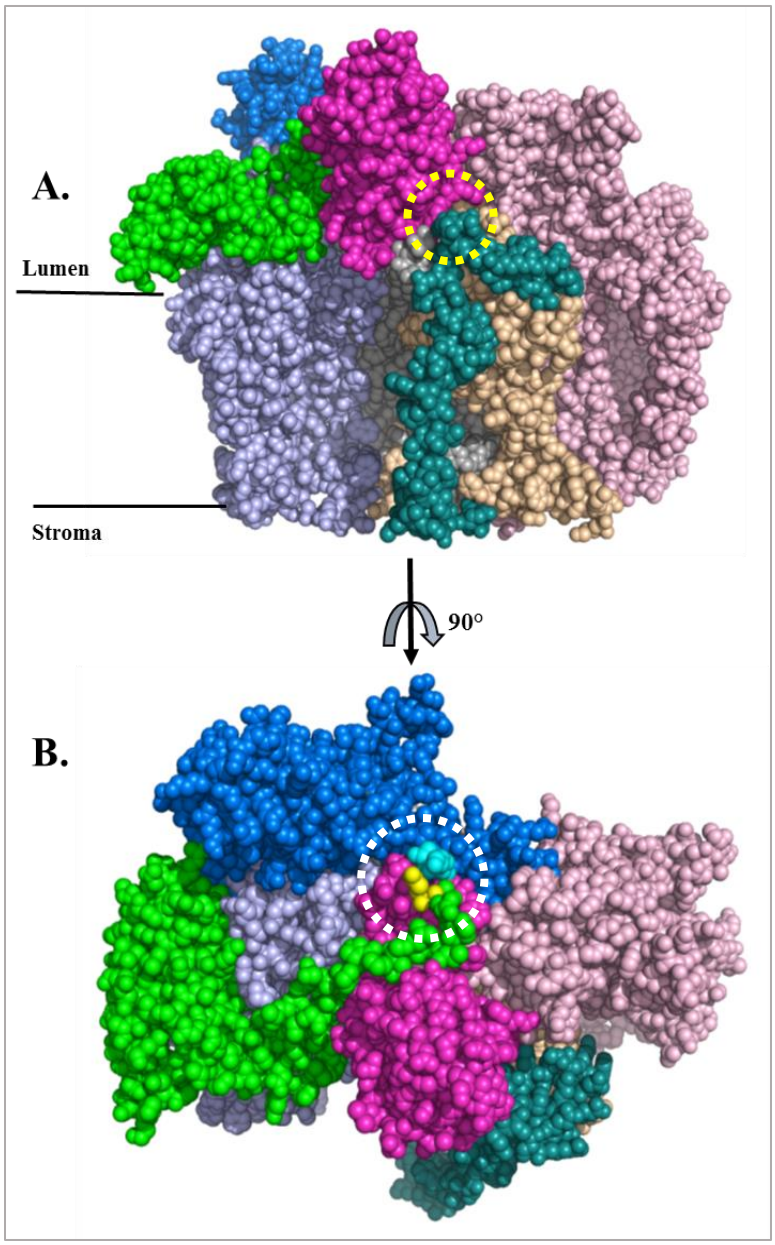


FIGURE 5.3 Cryo-Electron Microscopy Structure for spinach PSII-LHCII

Cryo-EM structure (Wei et al., 2016) for PSII core with extrinsic proteins PsbO, PsbP and PsbQ is illustrated here. **A.** PsbQ (green) identified with between PsbO (blue) and CP43 (light blue). PsbP (magenta) close to PsbE and PsbO. N-terminus of PsbP is close to PsbE: ^{57}E, shown with a dotted yellow circle **B.** 90° rotation of the above figure, N-terminus of PsbQ is an extended structure. PsbQ N-terminus extends up to PsbP. Crosslinked residues of PsbP and PsbQ are represented in cyan and yellow spheres respectively, and highlighted with a white dotted circle.

PsbO, PsbP and PsbQ, specifically seemed to shield the lumenal domains of CP43 and D1 (C-terminal domain). The location of PsbQ is similar to red alagal structure, where PsbQ is located between CP43 and PsbO proteins. The location of PsbO is similar to cyanobacteria and red algae, however, PsbQ is located closer to the N-terminus of PsbO and CP43 similar to the PsbQ' location in red algae. The PsbQ, unlike X-ray crystal structure, has an extended N-terminus that stretches from its core structure (helix-bundle) towards PsbP. Such, extended N-terminus for PsbQ was presented earlier by NMR studies (Rathner et al., 2015). The cryo-EM location of extrinsic proteins is shown in FIGURE 5.3 A and B. The PsbQ: ^{1}E is in van der Waals contact to ^{93}Y, ^{96}K and ^{97}T of PsbP (as seen in BS3 crosslinking results). Additionally, the structure of the N-terminus of PsbO is similar too. These studies not only confirm our models based on protein crosslinking data, but also suggest how protein crosslinking is advantageous in studying protein complexes.

References

Ago, H., Adachi, H., Umena, Y., Tashiro, T., Kawakami, K., Kamiya, N., Tian, L., Han, G., Kuang, T., Liu, Z., *et al.* (2016). Novel features of eukaryotic photosystem II revealed by its crystal structure analysis from a red alga. The Journal of biological chemistry.
Balsera, M., Arellano, J.B., Revuelta, J.L., de las Rivas, J., and Hermoso, J.A. (2005). The 1.49 A resolution crystal structure of PsbQ from photosystem II of Spinacia oleracea reveals a PPII structure in the N-terminal region. Journal of molecular biology *350*, 1051-1060.

Cao, P., Xie, Y., Li, M., Pan, X., Zhang, H., Zhao, X., Su, X., Cheng, T., and Chang, W. (2015). Crystal structure analysis of extrinsic PsbP protein of photosystem II reveals a manganese-induced conformational change. Molecular plant *8*, 664-666.

Ido, K., Nield, J., Fukao, Y., Nishimura, T., Sato, F., and Ifuku, K. (2014). Cross-linking evidence for multiple interactions of the PsbP and PsbQ proteins in a higher plant photosystem II supercomplex. The Journal of biological chemistry *289*, 20150-20157.

Kakiuchi, S., Uno, C., Ido, K., Nishimura, T., Noguchi, T., Ifuku, K., and Sato, F. (2012). The PsbQ protein stabilizes the functional binding of the PsbP protein to photosystem II in higher plants. Biochimica et biophysica acta *1817*, 1346-1351.

Liu, H., Zhang, H., Weisza, D.A., Vidavsky, I., Gross, M.L., and Pakrasi, H.B. (2014). MS-based cross-linking analysis reveals the location of the PsbQ protein in cyanobacterial photosystem II. PNAS *111*, 4638-4643.
Roose, J.L., Frankel, L.K., Mummadisetti, M.P., and Bricker, T.M. (2016). The extrinsic proteins of photosystem II: update. Planta.

Tomita, M., Ifuku, K., Sato, F., and Noguchi, T. (2009). FTIR evidence that the PsbP extrinsic protein induces protein conformational changes around the oxygen-evolving Mn cluster in photosystem II. Biochemistry *48*, 6318-6325.

Wei, X., Su, X., Cao, P., Liu, X., Chang, W., Li, M., Zhang, X. and Liu, Z. (2016). Structure of spinach photosystem II–LHCII supercomplex at 3.2 Å resolution. Nature *534*, 69-74.

APPENDIX A

Modelling of PsbP N-terminus

We used a program MODELLER to add secondary structure to the crystallographically unresolved N-terminus of PsbP. This process required three files for the N-terminal model building process. (1) An alignment file for the template sequence and target sequence. (2) Structure file (.pdb) for the template protein. (3) A script file, that assigns the tasks to the program MODELLER. The generated models were validated for their secondary structures by Ramachandran analysis using the program PROCHECK, and analyzed further for their z-scores with the program PROSA-II. The scripts were similar for modelling the N-terminus for a helix or random coil in both PsbP and PsbO, and hence only PsbP is shown below.

Alignment file used to model PsbP N-terminus

```
>P1;Manju_4RTI_modified.pdb
structureX:Manju_4RTI_modified.pdb:1:A:175:A:::0.00:0.00
----------PKKNTEFMPYNGDGFKLLVPSKWNPSKEKEFPGQVLRYEDNFDATSNLS
VLVQPTDKKSITDFGSPEDFLSQVDYLLGKQAYFGKTDSEGGFDSGVVASANVLESSTPV
VDGKQYYSITVLTRTADGDEGGKHQVIAATVKDGKLYICKAQAGDKRWFKGAKKFVESAT
SSFSVA*

>P1;manju_model1
sequence:manju_model1:1::186::::0.00:0.00
AYGEAANVFGKPKKNTEFMPYNGDGFKLLVPSKWNPSKEKEFPGQVLRYEDNFDATSNLS
VLVQPTDKKSITDFGSPEDFLSQVDYLLGKQAYFGKTDSEGGFDSGVVASANVLESSTPV
VDGKQYYSITVLTRTADGDEGGKHQVIAATVKDGKLYICKAQAGDKRWFKGAKKFVESAT
SSFSVA*
```

(1) Structure file

This file can be obtained from www.rscb.org. The program BLAST is generally used to identify the most appropriate target protein to be used as a template. Since spinach PsbP crystal structure is available, it was used as a template for this project. The gaps in the .pdb file were filled by

renumbering them, and the gaps were shown as dashes in the alignment file to represent the regions unresolved in the crystal structure.

(2) Script file (.py) to fill the gaps with loops, without adding restraints

```
# Loop PsbP

from modeller import *
from modeller.automodel import *

log.verbose()
env = environ()

class MyModel(automodel):
   def special_restraints(self, aln):
      rsr = self.restraints
      at = self.atoms


a = MyModel(env,
          alnfile='manju_4RTI.ali',  # alignment filename
          knowns=('Manju_4RTI_modified.pdb'), # codes of the templates
          sequence='manju_model1',    # code of the target
          assess_methods= (assess.DOPE,assess.GA341))  # request GA341 assessment
a.starting_model= 1                 # index of the first model
a.ending_model  = 20                  # index of the last model
                            # (determines how many models to calculate)


# Very thorough VTFM optimization:
a.library_schedule = autosched.slow
a.max_var_iterations = 300

# Thorough MD optimization:
a.md_level = refine.slow

# Repeat the whole cycle 2 times and do not stop unless obj.func. > 1E6
a.repeat_optimization = 2
a.max_molpdf = 1e6

a.make()                          # do homology modeling
```

(3) Designing Helix at the N-terminus of PsbP

```
# Helix PsbP

from modeller import *
from modeller.automodel import *

log.verbose()
env = environ()

class MyModel(automodel):
   def special_restraints(self, aln):
      rsr = self.restraints
      at = self.atoms

      # Residues 1 through 15 should be an alpha helix:
      rsr.add(secondary_structure.alpha(self.residue_range('1:', '15:')))
          # Restrain the specified CA-CA distance to 12.5 angstroms (st. dev.=0.1)
      # Use an upper bound (right Gaussian) potential and X-Y distance group.
          rsr.add(forms.upper_bound(group=physical.xy_distance, feature=features.distance(at['N:1'],
at['NZ:170']), mean=11.4, stdev=0.1))
          rsr.add(forms.upper_bound(group=physical.xy_distance, feature=features.distance(at['N:1'],
at['NZ:173']), mean=11.4, stdev=0.1))
          rsr.add(forms.upper_bound(group=physical.xy_distance, feature=features.distance(at['OH:2'],
at['NZ:170']), mean=11.4, stdev=0.1))
          rsr.add(forms.upper_bound(group=physical.xy_distance, feature=features.distance(at['NZ:13'],
at['NZ:174']), mean=11.4, stdev=0.1))
          rsr.add(forms.upper_bound(group=physical.xy_distance, feature=features.distance(at['NZ:14'],
at['NZ:174']), mean=11.4, stdev=0.1))

a = MyModel(env,
          alnfile='manju_2vu4A.ali',  # alignment filename
          knowns=('Manju_4RTI_modified.pdb'), # codes of the templates
          sequence='manju_model1',    # code of the target
          assess_methods= (assess.DOPE,assess.GA341))  # request GA341 assessment
a.starting_model= 1                 # index of the first model
a.ending_model  = 1                 # index of the last model
                            # (determines how many models to calculate)

# Very thorough VTFM optimization:
a.library_schedule = autosched.slow
a.max_var_iterations = 300

# Thorough MD optimization:
a.md_level = refine.slow

# Repeat the whole cycle 2 times and do not stop unless obj.func. > 1E6
a.repeat_optimization = 2
a.max_molpdf = 1e6

a.make()                        # do homology modeling
```

APPENDIX B

PERMISSION TO REPRINT

APPENDIX C

PERMISSION TO REPRINT

From: PNAS Permissions <PNASPermissions@nas.edu>
Sent: Wednesday, July 13, 2016 3:35 PM
To: Manjula P Mummadisetti
Subject: RE: Permission to use my article as a Ph.D. thesis chapter

Dear Dr. Mummadisetti

Authors do not need to obtain permission for the following uses of material they have published in PNAS: (1) to use their original figures or tables in their future works; (2) to make copies of their papers for their own personal use, including classroom use, or for the personal use of colleagues, provided those copies are not for sale and are not distributed in a systematic way; (3) to include their papers as part of their dissertations; or (4) to use all or part of their articles in printed compilations of their own works.

Yes, you may include this email as a permission for your dissertation! I've included our standard approval language below, in case you need PNAS to explicitly grant permission:

Permission is granted for your use of the material as described in your message. Please cite the PNAS article in full when re-using its material (whether it's the whole article, or figures/tables/excerpts). Because this material published after 2008, a copyright note is not needed. There is no charge, either. Let us know if you have any questions.

Best regards,
Kay McLaughlin for
Diane Sullenberger
Executive Editor

RIGHTS AND PERMISSIONS LINK FOR PNAS

http://www.pnas.org/site/aboutpnas/rightperm.xhtml

VITA

Ms. Manjula Mummadisetti was born and raised in India. She earned her B.Sc. degree in Microbiology, Botany and Chemistry from Kasturba Gandhi College, M.Sc. in Biochemistry from Osmania University and M.S. in Molecular Science and Nanotechnology from Louisiana Tech University. Her strong desire to obtain a Ph.D. degree brought her to United States. Ms. Mummadisetti is a candidate to receive a Ph.D. in Biochemistry from Louisiana State University in Summer of 2016.